

Fr m:
Sent:
To:
Subject:

Li, Ruixiang
Wednesday, June 12, 2002 2:47 PM
STIC-Biotech/ChemLib
Sequence search of Application NO: 09/727,739

Please do a standard search SEQ ID NOs: 3, 15, and 18 against both the commercial and interference amino acid database.

Thank you very much!

Ruixiang Li
GAU 1646
CM1 10E18
Mail Box 10C01
306-0282

CRFG

Searcher: D. Schreiber
Phone: 308-4292
Location: CM1 6A03
Date Picked Up: 6/13
Date Completed: 6/19
Searcher Prep/Review: 7
Clerical: _____
Online time: 6

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: 3
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST(where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: Camphgen
WWW/Internet: _____
Other (specify): _____

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: June 13, 2002, 12:23:38 ; Search time 35,79 Seconds
(without alignments)
298.014 Million cell updates/sec

Title: US-09-727-739B-15
Perfect score: 597
Sequence: 1 MRVSIHICALALLGLALAIC.....PPRERKAGCKNFYWKGFSTC 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	476	79.7	115	2 I51064	somatostatin II pr
2	194	32.5	125	1 RIAFS2	somatostatin II pr
3	179.5	30.1	116	1 S20630	somatostatin precu
4	176.5	29.6	116	1 RIHUS1	somatostatin I pre
5	176.5	29.6	116	1 A28968	somatostatin I pre
6	174	29.1	115	2 JC6166	somatostatin-14 pr
7	163.5	27.4	116	1 RIRTS1	somatostatin precu
8	163.5	27.4	116	1 RIMSS1	somatostatin precu
9	162.5	27.2	116	1 RIBOS1	somatostatin-14 pr
10	156.5	26.2	114	1 RIIDS1	preprosomatostatin
11	156.5	26.2	114	2 I50798	somatostatin-25 -
12	147	24.6	25	2 A60842	somatostatin II pr
13	136	22.8	73	2 S00169	somatostatin I pre
14	136	22.8	92	1 RIFGS	somatostatin I pre
15	131.5	22.0	121	1 RIAFS1	somatostatin II pr
16	125.5	21.0	74	2 S00166	somatostatin-25 -
17	114	19.1	25	2 B60840	somatostatin-28 -
18	105	17.6	28	2 A61322	somatostatin-relat
19	104	17.4	34	2 A32271	somatostatin, panc
20	94	15.7	37	2 A32000	somatostatin-14 [P
21	91.5	15.3	103	2 JC6167	cortistatin precu
22	87	14.6	112	2 S67489	cortistatin-like p
23	80.5	13.5	105	2 JC5414	structural polypro
24	77.5	13.0	1236	1 VHWVWE	somatostatin - sli
25	77	12.9	14	2 C60414	somatostatin I - c
26	77	12.9	14	2 B60842	somatostatin I - E
27	77	12.9	14	2 A60840	somatostatin I - s
28	77	12.9	14	2 S00172	genome polyprotein
29	74	12.4	1241	2 S26373	

30	74	12.4	1242	2 S72350	structural polypro
31	74	12.4	1242	2 A56605	structural polypro
32	73.5	12.3	342	2 A44965	cysteine proteinas
33	72	12.1	14	2 A60622	somatostatin - spo
34	70.5	11.8	221	2 C34768	ORF2 protein - Orf
35	70	11.7	243	2 C75608	hypothetical prote
36	69.5	11.6	342	2 A45524	cysteine proteinas
37	69.5	11.6	1305	2 T31096	cyclin G-associate
38	69	11.6	341	2 B86269	F21F23.4 protein -
39	69	11.6	758	2 T51335	subtilisin-like pr
40	69	11.6	1205	1 A38943	nitric-oxide synth
41	68.5	11.5	435	2 A82554	conserved hypothet
42	67.5	11.3	678	2 AB2744	two component sens
43	67.5	11.3	861	2 A97525	cell cycle histidi
44	67	11.2	120	2 F83244	hypothetical prote
45	67	11.2	1013	2 T04740	hypothetical prote

ALIGNMENTS

RESULT 1
I51064
somatostatin II precursor - rainbow trout
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C;Accession: I51064
R;Moore, C.A.; Kittilson, J.D.; Dahl, S.K.; Sheridan, M.A.
Gen. Comp. Endocrinol. 98, 253-261, 1995
A;Title: Isolation and characterization of a cDNA encoding for preprosomatostatin co
A;Reference number: I51064; MUID:95354921
A;Accession: I51064
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-115 <MOO>
A;Cross-references: EMBL:U32471; NID:g975344; PIDN:AAC59695.1; PID:g975345
C;Superfamily: somatostatin

Query Match 79.7%; Score 476; DB 2; Length 115;
Best Local Similarity 83.1%; Pred. No. 1.4e-43;
Matches 98; Conservative 3; Mismatches 7; Indels 10; Gaps 4;

QY 1 MRVSIHICALALLGLALAICSGAASQPDLDLASRRLLQALAAALPHRSVGSERWRTFY 60
Db 1 MKVCRHICALALLGLALAICSGAASQPDLDLSRRLLQARAAAAPHRSVGSERWRTFY 60
QY 61 PNCPCLRWRPRKVKGPOLKAKED---LERSV---DNLPPEKAGCKNFYWKGFSTC 111
Db 61 PNCPCCL--RPRKVKCP-AGAKEDLRVELERSVGNPNLPPRERKAGCKNFYWKGFSTC 115

RESULT 2
RIAFS2

somatostatin II precursor - American goosefish
C;Species: Lopholagus americanus (American goosefish)
C;Date: 31-Mar-1981 #sequence_revision 31-Mar-1981 #text_change 28-May-1999
C;Accession: B93236; A94038; A27376; A01434; A21881; A93236
R;Hobart, P.; Crawford, R.; Shen, L.; Pictet, R.; Rutter, W.J.
Nature 288, 137-141, 1980
A;Title: Cloning and sequence analysis of cDNAs encoding two distinct somatostatin pr
A;Reference number: A93236; MUID:81052423
A;Accession: B93236
A;Molecule type: mRNA
A;Residues: 1-125 <HOB>
A;Cross-references: GB:V00641; GB:J00947; GB:M23199; NID:g64030; PIDN:CAA23987.1; PII
A;Experimental source: islet tissue (endocrine pancreas)
R;Spless, J.; Noe, B.D.
Proc. Natl. Acad. Sci. U.S.A. 82, 277-281, 1985
A;Title: Processing of an anglerfish somatostatin precursor to a hydroxylysine-contai
A;Reference number: A94038; MUID:85113184
A;Accession: A94038
A;Molecule type: protein

A28968
Somatostatin I precursor - crab-eating macaque
N:Alternate names: preprosomatostatin
N:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)
C:Species: Macaca fascicularis (crab-eating macaque)
C:Date: 30-Jun-1989 #sequence_revision 31-Jan-1997 #text_change 18-Jun-1999
C:Accession: A28968
R:Travis, G.H.; Sutcliffe, J.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 1696-1700, 1988
A:Title: Phenol emulsion-enhanced DNA-driven subtractive cDNA cloning: isolation of low-abundance rat somatostatin-14 cDNA
A:Reference number: A28968; MUID:88144503
A:Accession: A28968
A:Molecule type: mRNA
A:Residues: 1-116 <TRA>
A:Cross-references: GB:M19318; NID:g342298; PIDN:AAA36908.1; PID:g342299
C:Comment: Somatostatin inhibits the release of somatotropin.
C:Superfamily: somatostatin
C:Keywords: hormone; neuropeptide
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-88/Domain: propeptide #status predicted <PRO>
F:89-116/Product: somatostatin-28 #status predicted <M28>
F:103-116/Product: somatostatin-14 #status predicted <M14>
F:105-116/Disulfide bonds: #status predicted

Query Match 29.6% Score 176.5; DB 1; Length 116;
Best Local Similarity 38.2%; Pred. No. 1.3e-11;
Matches 47; Conservative 19; Mismatches 30; Indels 27; Gaps 6;

QY 5 QHICALALLGLALAI-CSQGAASQPDLDLASRRLLQRALAAALPHRSGVSRWTFYPNC 63
Db :: ||||| : ||: | || | | : ||: ||||| : | | : |
5 RLQCALAALSIVLALGCVTGAPSPRL----RQFLQKSLAA-----AGKQELAKYFLAE- 55

QY 64 PCLRWRPRKVKGPOLKAKE-----DLERSVDNLP---PRRKAGCKNFKY 108
Db :: ||||| : ||: | || | | : ||: ||||| : | | : |
56 --LLSEPNQTDNALEPEDLSQAABQDEMRLELORSANSNPAMAPRRKAGCKNFKWTF 113

QY 109 TSC 111
Db |||
114 TSC 116

RESULT 6
JC6166
Somatostatin-14 precursor - laughing frog
N:Alternate names: PSS1 protein
C:Species: Rana ridibunda (laughing frog)
C:Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 16-Jul-1999
C:Accession: JC6166
R:Tostivint, H.; Lihmann, I.; Bucharles, C.; Vieau, D.; Coulouarn, Y.; Fournier, A.; Co
Proc. Natl. Acad. Sci. U.S.A. 93, 12605-12610, 1996
A:Title: Occurrence of two somatostatin variants in the frog brain: Characterization of
A:Reference number: JC6166; MUID:97057290
A:Contents: brain
A:Accession: JC6166
A:Molecule type: mRNA
A:Residues: 1-115 <TOS>
A:Cross-references: GB:J068136; NID:gl890650; PIDN:AAC60093.1; PID:gl890651
C:Comment: This protein acts both as a neurotransmitter/neuromodulator and a hormone.
C:Genetics:
A:Gene: pss1
C:Superfamily: somatostatin
C:Keywords: brain; hormone

Query Match 29.1% Score 174; DB 2; Length 115;
Best Local Similarity 39.4%; Pred. No. 2.4e-11;
Matches 50; Conservative 17; Mismatches 32; Indels 28; Gaps 6;

QY 1 MRVSIHCALALLGLALAI-CSQGAASQPDLDLASRRLLQRALAAALPHRSGVSRWTF 59
Db :: ||||| : ||: | || | | : ||: ||||| : | | : |
1 MQSCRVCALTLLSLALAINSISAPTDRL----RQFLQKSLASA-----GKQELAKYF 51

QY 60 YPNCPLRWRPRKVKGPOLKAKE-----DLERSVDNLP---PRRKAGCKNFKY 104
Db :: ||||| : ||: | || | | : ||: ||||| : | | : |
52 LAE---LLSEPSQTDNEALESDDLPGAEQDEVRLERSANSPPALAPRRKAGCKNFF 108

QY 105 WKGTSC 111
Db || |||||
109 WKTTSC 115

RESULT 7
RIRTS1
Somatostatin precursor - rat
N:Alternate names: preprosomatostatin
N:Contains: somatostatin-14; somatostatin-28
C:Species: Rattus norvegicus (Norway rat)
C:Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 18-Jun-1999
C:Accession: A20983; A01431; A47598; A22529; I55220; I51829
R:Montminy, M.R.; Goodman, R.H.; Horovitch, S.J.; Habener, J.F.
Proc. Natl. Acad. Sci. U.S.A. 81, 3337-3340, 1984
A:Title: Primary structure of the gene encoding rat preprosomatostatin.
A:Reference number: A20983; MUID:84221954
A:Accession: A20983
A:Molecule type: DNA
A:Residues: 1-116 <MON>
A:Cross-references: GB:J00787; NID:g207024; PIDN:AAA42164.1; PID:g207025
A:Note: the authors translated the codon ACC for residue 43 as Tyr
R:Argos, P.; Taylor, W.L.; Minth, C.D.; Dixon, J.E.
J. Biol. Chem. 258, 8788-8793, 1983
A:Title: Nucleotide and amino acid sequence comparisons of preprosomatostatins.
A:Reference number: A01431; MUID:83238516
A:Accession: A01431
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-116 <ARG>
R:Benoit, R.; Ling, N.; Esch, F.
Science 238, 1126-1129, 1987
A:Title: A new prosomatostatin-derived peptide reveals a pattern for prohormone clea
A:Reference number: A47598; MUID:88070564
A:Accession: A47598
A:Status: preliminary
A:Molecule type: protein
A:Residues: 25-34 <BEN>
R:Tavianini, M.A.; Hayes, T.E.; Magazin, M.D.; Minth, C.D.; Dixon, J.E.
J. Biol. Chem. 259, 11798-11803, 1984
A:Title: Isolation, characterization, and DNA sequence of the rat somatostatin gene.
A:Reference number: A22529; MUID:85006903
A:Accession: A22529
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-116 <TAV>
A:Cross-references: GB:K02248; NID:g207014; PIDN:AAA42161.1; PID:g207017
R:Goodman, R.H.; Jacobs, J.W.; Dee, P.C.; Habener, J.F.
J. Biol. Chem. 257, 1156-1159, 1982
A:Title: Somatostatin-28 encoded in a cloned cDNA obtained from a rat medullary thyr
A:Reference number: I55220; MUID:82120034
A:Accession: I55220
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 38-78, 'H', '80-116 <RES>
A:Cross-references: GB:J00788; NID:g207018; PIDN:AAA42162.1; PID:g207019
R:Goodman, R.H.; Montminy, M.R.; Low, M.J.; Habener, J.F.
Adv. Exp. Med. Biol. 188, 31-47, 1985
A:Title: Biosynthesis of rat preprosomatostatin.
A:Reference number: I51829; MUID:85303584
A:Accession: I51829
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-116 <RE2>
A:Cross-references: GB:M25890; NID:g207030; PIDN:AAA42167.1; PID:g207031
C:Comment: Somatostatins are found in a variety of tissues, including hypothalamus,
C:Genetics:
A:Introns: 46/3
C:Function:

A:Molecule type: protein
A:Residues: 101-114 <AND>
C:Superfamily: somatostatin
C:Keywords: neuropeptide
F:1-24/Domain: signal sequence
F:25-100/Domain: propeptide
F:101-114/Product: somatostatin
F:103-114/Disulfide bonds: 1

Query Match	26.2%	Score 156.5;	DB 1;	Length 114;
Best Local Similarity	31.4%	Pred. No. 1.7e-09;		
Matches 43;	Conservative	20;	Mismatches 25;	Indels 49;
			Gaps	5;

QY
60 YPNCPCCLRWRPRKVKGPOLKAK-----EDLERSVDN-----LPPR 94
| : : : | : : : : :
42 ---PSVKGELTRYTLAEILAELOAOAENEVLDSDVSRAAESGARLEMERAAAGPMIAPR 97

```
QY 95 ERKAGCKNFYWKGETSC 111
      |||||:| ||||
DB 98 ERKAGCKNEFWKTFETSC 114
```

RESULT 11
I50798
preprosomatostatin SS-14 - channel catfish
C:Species: Ictalurus punctatus (channel catfish)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: I50798
R:Dixon, J.E.; Andrews, P.C.
Adv. Exp. Med. Biol. 188, 19-29, 1985
A:Title: Somatostatins of the channel catfish.
A:Reference number: I50798; MUID:85303576
A:Accession: I50798
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-114 <DIX>
A:Cross-references: GB:M25903; NID:g213339; PIDN:AAA49339.1; PID:g213340
C:Superfamily: somatostatin

Query Match	26.2%;	Score 156.5;	DB 2;	Length 114;
Best Local Similarity	31.4%;	Pred. NO. 1.7e-09;		
Matches 43; Conservative	20;	Mismatches 25;	Indels 49;	Gaps 5;

QY 60 YPNCPCLLRRPRKVKGPQLKAK-----EDLRSVDN-----LPFR 94
| :: : : | : : : : : : : :
Db 42 ----PSVROELTRYTLAEILAEIAFAENEVIQSDEVSAAASEGARYEMERPAAGDMADP 97

QY 95 ERKAGCKNFYWKGTSC 111
 '|||||:| | | |
 DQ 98 ERKAGCKNEFWKFTSC 114

RESULT 12
A60842
somatostatatin-25 - coho salmon
N:Alternate names: somatostatin II precursor
C:Species: Oncorhynchus kisutch (coho salmon)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: A60842; C60842
R:Plietskaya, E.M.; Pollock, H.G.; Rouse, J.B.; Hamilton, J.W.; Kimmel, J.R.; Andrews,
Gen. Comp. Endocrinol. 63, 252-263, 1986
A:Title: Characterization of Coho salmon (Oncorhynchus kisutch) islet somatostatins.
A:Reference number: A60842; MUID:87055212

A:Accession: A60842
A:Molecule type: protein
A:Residues: 1-25 <P1>
A:Accession: C60842
A:Molecule type: protein
A:Residues: 12-25 <P12>
A>Note: this form, somatostatin
C:Superfamily: somatostatin
C:Keywords: hormone; pan
F:1-25/Product: somatostatin
F:12-25/Product: somatostatin

Query Match	24.6%;	Score 147;	DB 2;	Length 25;
Best Local Similarity	100.0%;	Pred. No. 3.8e-09;		
Matches 25: Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

RESULT 13
S00169

S00169
somatostatin II precursor - European flounder (tentative sequence) (fragments)
C:Species: *Platichthys flesus* (European flounder)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 31-Mar-2000
C:Accession: S00169
R:Conlon, J.M.; Davis, M.S.; Falkmer, S.; Thim, L.
Eur. J. Biochem. 168, 647-652, 1987
A:Title: Structural characterization of peptides derived from prosomatostatins
A:Reference number: S00166: MUID:88029486

A:Accession: S00109
 A:Molecule type: protein
 A:Residues: 1-10;11-45;46-73 <CON>
 A:Note: three peptides which probably originate from a common precursor, were isolated
 C:Superfamily: somatostatin
 C:Keywords: glycoprotein; neuropeptide; pancreatic islet
 F:1-10/Product: peptide F1 #status experimental <PF1>
 F:11-45/Product: peptide F3 #status experimental <PF3>
 F:46-73/Product: peptide F2 #status experimental <PF2>
 F:52-73/Disulfide bonds: #status experimental

Query Match 22.8%; Score 136; DB 2; Length 73;
Best Local Similarity 51.9%; Pred. No. 1.7e-07;
Matches 27; Conservative 4; Mismatches 5; Indels 16; Gaps 2;

RESULT 14

somatostatin I precursor - pig (fragment)
 N:Alternate names: prosomatostatin
 N:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)
 C:Species: *Sus scrofa domestica* (domestic pig)
 C:Date: 30-Nov-1980 #sequence revision 31-Jan-1997 #text_change 31-Jan-1997
 C:Accession: A34109; A24222; A91273; A93054; A90398; S13616; A01432
 R:Bersani, M.; Thim, L.; Baldissiera, F.G.A.; Holst, J.J.
 J. Biol. Chem. 264, 10633-10636, 1989
 A:Title: Prosomatostatin 1-64 is a major product of somatostatin gene expression
 A:Reference number: A34109; MUID:89278131

A:Reference number: A34109, MUID:89276151
A:Accession: A34109
A:Molecule type: protein
A:Residues: 1-64 <BER>
R:Schmidt, W.E.; Mutt, V.; Kratzin, H.; Carlquist, M.; Conlon, J.M.; Creutzfeldt, W.
FEBS Lett. 192, 141-146, 1985
A:Title: Isolation and characterization of prossl-32, a peptide derived from the N-t
A:Reference number: A24222; MUID:86030691
A:Accession: A24222

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:31:02 ; Search time 60.72 Seconds
(without alignments)
316.246 Million cell updates/sec

Title: US-09-727-739B-15
Perfect score: 597
Sequence: 1 MRYSQIHICALALLGLALAIC.....PPRRKAGCKNFYWKGETSC 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

total number of hits satisfying chosen parameters: 562222

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : SPREMBL_19:*
1:  sp_archaea:*
2:  sp_bacteria:*
3:  sp_fungi:*
4:  sp_human:*
5:  sp_invertebrate:*
6:  sp_mammal:*
7:  sp_mhc:*
8:  sp_organelle:*
9:  sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_bacheap:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	239.5	40.1		120	13	Q90Y39	Q90Y39 catostomus
2	236	39.5		115	13	Q90Y43	Q90Y43 osteoglossu
3	228.5	38.3		114	13	Q90Y42	Q90Y42 pantodon bu
4	192.5	32.2		114	13	Q90Y41	Q90Y41 gnathonemus
5	178.5	29.9		114	13	Q90Y40	Q90Y40 chitala chi
6	174.5	29.2		116	13	Q90XE1	Q90xe1 acipenser t
7	123	20.6		28	13	Q9PRN9	Q9prn9 carassius a
8	114	19.1		25	13	Q9PRV0	Q9prv0 anguilla ja
9	104	17.4		107	13	Q9DDE4	Q9dde4 brachydanio
10	96.5	16.2		111	13	Q90XE0	Q90xe0 acipenser t
11	85.5	14.3		23	13	Q9PRV6	Q9prv6 anguilla ja
12	79.5	13.3		105	11	Q9R1P8	Q9rlp8 mus musculu
13	77	12.9		1242	12	Q9PZX3	Q9pzx3 eastern equ
14	76	12.7		251	16	Q98HG9	Q98hg9 rhizobium l
15	75.5	12.6		259	12	Q86972	Q86972 western equ
16	75.5	12.6		1235	12	Q9IBP3	Q9ibp3 western equ

17	75.5	12.6	1236	12	Q9J1K1	Q9J1K1 western equ
18	75	12.6	1242	12	Q9PZX7	Q9PZX7 eastern equ
19	75	12.6	1242	12	Q9PZX6	Q9PZX6 eastern equ
20	74	12.4	1241	12	Q66579	Q66579 eastern equ
21	74	12.4	1242	12	Q9PZX5	Q9PZX5 eastern equ
22	74	12.4	1242	12	Q9PZX4	Q9PZX4 eastern equ
23	74	12.4	1242	12	Q9PZX2	Q9PZX2 eastern equ
24	74	12.4	1242	12	Q9PZX1	Q9PZX1 eastern equ
25	74	12.4	1242	12	Q9PZX0	Q9PZX0 eastern equ
26	74	12.4	1242	12	Q9PZW9	Q9PZW9 eastern equ
27	74	12.4	1242	12	Q9PZW8	Q9PZW8 eastern equ
28	74	12.4	1242	12	Q9PZW7	Q9PZW7 eastern equ
29	74	12.4	1242	12	Q9PZW6	Q9PZW6 eastern equ
30	74	12.4	1242	12	Q88790	Q88790 eastern equ
31	74	12.4	1242	12	Q88792	Q88792 eastern equ
32	74	12.4	1242	12	Q88793	Q88793 eastern equ
33	74	12.4	1242	12	Q88794	Q88794 eastern equ
34	74	12.4	1242	12	Q88795	Q88795 eastern equ
35	74	12.4	1242	12	Q88796	Q88796 eastern equ
36	74	12.4	1242	12	Q88797	Q88797 eastern equ
37	74	12.4	1242	12	Q88798	Q88798 eastern equ
38	74	12.4	1242	12	Q88799	Q88799 eastern equ
39	74	12.4	1242	12	Q08359	Q08359 eastern equ
40	72	12.1	210	4	Q9BSE9	Q9BSE9 homo sapien
41	71.5	12.0	1255	12	Q92604	Q92604 rice ragged
42	71	11.9	638	11	Q9D2N8	Q9D2N8 mus musculus
43	70.5	11.8	221	12	Q85299	Q85299 orf virus
44	70.5	11.8	571	4	Q15913	Q15913 homo sapien
45	70	11.7	243	16	Q9R3Z5	Q9R3Z5 deinococcus

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	120 AA.
Q90Y39				
ID	Q90Y39			
AC	Q90Y39;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	PREPROSMATOSTATIN.			
OS	Catostomus commersoni (White sucker).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostei; Teleostei; Euteleostei; Osteichthyes; Cypriniformes; Catostomidae; Catostomus.			
OC	Actinopterygii; Neopterygii;			
OC	Cypriniformes; Catostomidae; Catostomus.			
OX	NCBI_TaxID=7971;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Al-Mahrouki A.A., Irwin D.M., Youson J.H.;			
RT	"Molecular cloning and characterization of white sucker preprosmatostatin.";			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF292654; AAK97071.1;			
DR	EMBL; 120 AA; 13783 MW;			
DR	00828D35263E8805 CRC64;			

```
Query Match          40.1%; Score 239.5; DB 13; Length 120;
Best Local Similarity 40.1%; Pred. No. 3.9e-19;
Matches 55; Conservative 18; Mismatches 21; Indels 43; Gaps 5;
```

OY 1 MRYSQIHCAALLGLALAI CSQAASQ--PDLDIASRRLQLRALAALPHRSVGVSERWRT 58
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 1 MRCELHCYLALLGSLVLC DRCADSQLPDMDFHRRLQLQRARAIGLATQD----- 52

OY 59.FYPNCPLRRWRPRKVG-----POLKAKE-----DLERSVDN---LPPR 94
- - - - - : : : : : | | | | | : | | | | :

Dd 53 -----WTKKDIEELLSQLSPLEIAREANGSVTTGGNDLLHLELSAESANTQLPYR 103


```

RESULT 2
Q90Y43
ID Q90Y43 PRELIMINARY; PRT; 115 AA.
AC Q90Y43;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Osteoglossum bicirrhosum (silver arawana).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
OX NCBI_TaxID=109271;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292650; AAK97067.1;
SQ SEQUENCE 115 AA; 12791 MW; D65FBD7C6F1E4E4D CRC64;

Query Match 39.5%; Score 236; DB 13; Length 115;
Best Local Similarity 47.5%; Pred. No. 9.3e-19;
Matches 58; Conservative 10; Mismatches 36; Indels 18; Gaps 4;

QY 1 MRVSIHCALALLGLALAIICSGAASQPDLDLASRRLLQRLAALPHRSGVSRWRTFY 60
Db 1 MKICQIHCTLVLLGLVGLYCPSAASQP--DLRYRSFLQRAHAAMS-----PDNWSKQA 53

QY 61 PNCPCLRWRPRKVKPQ-----LKAKEDLERSVD--NLPPRERKAGCKNFYWKGFT 109
Db 54 VEELLSRLAPAGQEVPPQGAASQPDLDLASRRLLQRLAALPHRSGVSRWRTFY 113

QY 110 SC 111
Db 114 SC 115

RESULT 3
Q90Y42
ID Q90Y42 PRELIMINARY; PRT; 114 AA.
AC Q90Y42;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Pantodon buchholzi (Butterflyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Pantodontidae; Pantodon.
OX NCBI_TaxID=8276;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292651; AAK97068.1;
SQ SEQUENCE 114 AA; 12352 MW; 7E3D44CB6A27B12F CRC64;

Query Match 38.3%; Score 228.5; DB 13; Length 114;
Best Local Similarity 41.4%; Pred. No. 6.3e-18;
Matches 55; Conservative 14; Mismatches 23; Indels 41; Gaps 5;

QY 1 MRVSIHCALALLGLALAIICSGAASQPDLDLASRRLLQRLAALPHRSGVSRWRTFY 60
Db 1 MKLCQVHCILALLGLVLMCGSSSATQ--LDSRYRSLVQRAAASMGPOD----- 48

```

```

QY 61 PNCPCLRWRPRKVK-----GPQLKAKE-----DLERSVD--NLPPRERKA 98
Db 49 -----WGKLSVEDLSLLAATEADMFQDMSAAESEGAHLDLERSVEPGNVPPRKA 101

QY 99 GCKNFYWKGTSC 111
Db 102 GCKNFYWKGTSC 114

RESULT 4
Q90Y41
ID Q90Y41 PRELIMINARY; PRT; 114 AA.
AC Q90Y41;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Gnathonemus petersii.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Gnathonemus.
OX NCBI_TaxID=42645;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292652; AAK97069.1;
SQ SEQUENCE 114 AA; 12494 MW; 454DA57A309CA8F2 CRC64;

Query Match 32.2%; Score 192.5; DB 13; Length 114;
Best Local Similarity 41.4%; Pred. No. 6.5e-14;
Matches 55; Conservative 8; Mismatches 29; Indels 41; Gaps 5;

QY 1 MRVSIHCALALLGLALAIICSGAASQPDLDLASRRLLQRLAAL--PHRSGV 52
Db 1 MLSSRIQCALALLSLALAVSSVSAAPS---DLKRLQLQRLSLAPASKQDLARNPLELL 57

QY 53 SERWRTFVPCPCLRWRPRKVKGPOLKAKEDLERSVDN-----LPPRERKA 98
Db 58 SEMVR-----VENEALPDDLRSRGADQEEVRLERAAAGPALAPRERKA 101

QY 99 GCKNFYWKGTSC 111
Db 102 GCKNFYWKGTSC 114

RESULT 5
Q90Y40
ID Q90Y40 PRELIMINARY; PRT; 114 AA.
AC Q90Y40;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Chitala chitala.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Notopteridae; Chitala.
OX NCBI_TaxID=112163;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292653; AAK97070.1;
SQ SEQUENCE 114 AA; 12561 MW; 4E3C32F58E34F971 CRC64;

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:22:20 ; Search time 72.45 Seconds
(without alignments)
170.175 Million cell updates/sec

Title: US-09-727-739B-15
Perfect score: 597
Sequence: 1 MRVSIHICALALLGLALAIC.....PPRERKAGCKNFYWKGTSC 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues 747574
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	597	100.0	111	22 AAU07668	Rainbow trout prep
2	482	80.7	115	22 AAU07667	Rainbow trout prep
3	194	32.5	125	3 AAP20029	Sequence of prepro
4	176.5	29.6	116	21 AAG03774	Human secreted pro
5	164.5	27.6	114	22 AAU07666	Rainbow trout prep
6	132	22.1	28	7 AAP61714	Somatostatin-28 an
7	132	22.1	28	10 AAP90989	Analogue of angler
8	131.5	22.0	121	3 AAP20028	Sequence of prepro
9	105.5	17.7	140	21 AAY69789	MWpsp-MWpmp20-(His
10	105	17.6	25	3 AAP20198	Sequence of somato
11	105	17.6	25	22 AAB91017	Somatostatin relat

12	105	17.6	28	19 AAW51859	Somatostatin analo
13	105	17.6	28	20 AAY28703	Mouse somatostatin
14	105	17.6	28	20 AAY24384	Somatostatin pepti
15	105	17.6	28	20 AAY24239	Peptide hormone so
16	105	17.6	28	22 AAB91018	Somatostatin relat
17	105	17.6	28	22 AAU07669	Mammalian somatost
18	105	17.6	29	22 AAB91020	Somatostatin relat
19	104	17.4	20	6 AAP50414	Swine duodenum eic
20	104	17.4	27	22 AAB91007	Somatostatin relat
21	104	17.4	28	19 AAW50950	Somatostatin analo
22	104	17.4	28	22 AAB91021	Somatostatin relat
23	101	16.9	28	19 AAW50816	Somatostatin-28 an
24	100	16.8	28	3 AAP20130	Somatostatin-28 pe
25	100	16.8	28	3 AAP20131	Somatostatin-28 pe
26	100	16.8	28	19 AAW50818	Somatostatin-28 an
27	98	16.4	28	3 AAP20197	Sequence of [D-Trp
28	97	16.2	28	3 AAP20125	Somatostatin-28 pe
29	97	16.2	28	19 AAW50819	Somatostatin analo
30	97	16.2	28	19 AAW51855	Somatostatin analo
31	97	16.2	28	19 AAW51688	Somatostatin analo
32	96	16.1	28	4 AAP30058	Radioactively tagg
33	92	15.4	28	19 AAW50817	Somatostatin-28 an
34	91	15.2	28	3 AAP20126	Somatostatin-28 pe
35	89	14.9	14	22 AAU07665	Rainbow trout soma
36	89	14.9	28	3 AAP20128	Somatostatin-28 pe
37	87	14.6	112	19 AAW42033	Rat preprocartista
38	84	14.1	189	22 ABG22878	Novel human diagno
39	83	13.9	14	7 AAP60190	Somatostatin-28 an
40	83	13.9	36	21 AAY58743	Somatostatin-28 an
41	83	13.9	54	21 AAY58760	Somatostatin-28 an
42	83	13.9	85	19 AAW42035	Somatostatin-28 an
43	80.5	13.5	105	19 AAW64433	Human cortistatin
44	80.5	13.5	105	19 AAW44045	HCS-105 peptide.
45	80.5	13.5	105	19 AAW42047	Human preprocortis

ALIGNMENTS

RESULT 1
AAU07668
ID AAU07668 standard; Protein; 111 AA.
XX AAU07668;
AC AAU07668;
XX 04-DEC-2001 (first entry)
DT Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.
DE Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.
XX Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
KW PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;
KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;
KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;
KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.
XX Oncochrychus mykiss.

Key	Location/Qualifiers
Peptide	1..25
Protein	/note= "Signal peptide"
Protein	/note= "PPSS-II' pre-sequence"
Protein	26..111
Peptide	/note= "Mature PPSS-II'"
Peptide	/note= "PPSS-II' pro-sequence"
Peptide	87..111
Cleavage-site	/note= "Prosomatostatin II'"
	96..97
	/note= "Dibasic cleavage site"

FT Peptide 98..111 /note= "SS-14 variant peptide"

FT CA2325169-A1.

PN 03-JUN-2001.

XX 01-DEC-2000; 2000CA-2325169.

XX 03-DEC-1999; 99US-0168934.

XX (NDSU-) NDSU RES FOUND.

PA Sheridan MA, Moore CA, Kittelson JD;

PI WPI; 2001-425997/46.

XX N-PSDB; AAS12935.

DR New somatostatin polypeptides derived from Oncorhynchus mykiss, useful for treating diabetes mellitus, acromegaly, gastrinoma, acquired immunodeficiency syndrome and neurological disorders

XX Claim 1; Fig 3; 52pp; English.

XX The invention relates to an Oncorhynchus mykiss somatostatin polypeptide containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of preprosomatostatin II (PPSS-II). The protein sequences and their associated polynucleotides are useful for identifying modified somatostatin polypeptide which functions as a somatostatin agonist useful for research, therapeutics or diagnostics, including medical and veterinary applications. The wild-type somatostatin and its modified version are useful for treating hypersecretion from endocrine tumours in the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g. gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage through their effects on cell proliferation and apoptosis and as adjuncts in the treatment of diabetes mellitus via inhibition of growth hormone and glucagon. In addition, dysfunctional somatostatin secretion is associated with acquired immunodeficiency syndrome (AIDS) and various neurological disorders (e.g. epilepsy, Alzheimer's disease and Huntington's disease) and somatostatin antagonists are effective in the treatment of such conditions. Nucleic acids encoding the polypeptides are useful in gene therapy and fusion peptides can be targeted to neoplasms and their metastases, inhibiting the release of their secretory products. This sequence represents O. Mykiss PPSS-II', protein.

XX Note: The features for this sequence are specifically claimed in the specification.

Sequence 111 AA;

Query Match 100.0%; Score 597; DB 22; Length 111;

Best Local Similarity 100.0%; Pred. No. 1.6e-63;

Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRVSIQHCAALLGLALAIQSGAASQPDLDASRRLLQRALAALPHRSGVSEWRRTFY 60

Db 1 mrvsqihcalallglalalcsggaasqpdldasrrllqralaaalphrsgvserwrtfy 60

Qy 61 PNCPCLRWRPRKVKGPQLKAKEDLERSVDNLPERRKAGCKNFYWKGTSC 111

Db 61 pncpcrlwrprkvkgpqlkakedlrsvdnlpprrkagcknfykgtsc 111

RESULT 2

AAU07667

XX AAU07667 standard; Protein; 115 AA.

XX AAU07667;

XX 04-DEC-2001 (first entry)

XX Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.

KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I; PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS; gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus; carcinoid syndrome; cell proliferation; apoptosis; growth hormone; glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV; epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective; neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic; anti-human immunodeficiency virus; osteopathic; anticonvulsant.

XX Oncorhynchus mykiss.

XX Key Location/Qualifiers

FT Peptide 1..25 /note= "Signal peptide"

FT Protein 1..87 /note= "PPSS-II' pre-sequence"

FT Protein 26..115 /note= "Mature PPSS-II'"

FT Misc-difference 74 /note= "Encoded by CAA"

FT Peptide 88..101 /note= "PPSS-II' pro-sequence"

FT Peptide 88..115 /note= "Prosomatostatin II'"

FT Cleavage-site 100..101 /note= "Dibasic cleavage site"

FT Peptide 102..115 /note= "SS-14 variant peptide"

XX CA2325169-A1.

XX 03-JUN-2001.

XX 01-DEC-2000; 2000CA-2325169.

XX 03-DEC-1999; 99US-0168934.

XX (NDSU-) NDSU RES FOUND.

XX Sheridan MA, Moore CA, Kittelson JD;

XX WPI; 2001-425997/46.

XX N-PSDB; AAS12934.

XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful for treating diabetes mellitus, acromegaly, gastrinoma, acquired immunodeficiency syndrome and neurological disorders

XX Claim 2; Fig 3; 52pp; English.

XX The invention relates to an Oncorhynchus mykiss somatostatin polypeptide containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of preprosomatostatin II (PPSS-II). The protein sequences and their associated polynucleotides are useful for identifying modified somatostatin polypeptide which functions as a somatostatin agonist useful for research, therapeutics or diagnostics, including medical and veterinary applications. The wild-type somatostatin and its modified version are useful for treating hypersecretion from endocrine tumours in the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g. gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage through their effects on cell proliferation and apoptosis and as adjuncts in the treatment of diabetes mellitus via inhibition of growth hormone and glucagon. In addition, dysfunctional somatostatin secretion is associated with acquired immunodeficiency syndrome (AIDS) and various neurological disorders (e.g. epilepsy, Alzheimer's disease and Huntington's disease) and somatostatin antagonists are effective in the treatment of such conditions. Nucleic acids encoding the polypeptides are useful in gene therapy and fusion peptides can be targeted to neoplasms and their metastases, inhibiting the release of their secretory products. This sequence represents O. Mykiss PPSS-II' protein.

XX Note: The features for this sequence are specifically claimed in the specification.

Db 56 --llsepnqtendalepedisqaaeqdemrlqlrsansnpamaprerkagcknffwktf 113
QY 109 TSC 111
Db 114 tsc 116
RESULT 5
AAU07666
ID AAU07666 standard; Protein; 114 AA.
XX AAU07666;
AC
DT 04-DEC-2001 (first entry)
XX
DE Rainbow trout preprosomatostatin I (PPSS-I) polypeptide.
XX
KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
cardinoid syndrome; cell proliferation; apoptosis; growth hormone;
glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;
neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;
anti-human immunodeficiency virus; osteopathic; anticonvulsant.
OS Oncoerhynchus mykiss.
XX
FH Key Location/Qualifiers
FT Peptide 1..24 /note= "Signal peptide"
FT Protein 1..88 /note= "PPSS-I pre-sequence"
FT Protein 25..114 /note= "Mature PPSS-I"
FT Peptide 89..100 /note= "PPSS-I pro-sequence"
FT Peptide 89..114 /note= "Prosomatostatin I"
FT Cleavage-site 99..100 /note= "Dibasic cleavage site"
FT Peptide 101..114 /note= "SS-14 peptide"
XX CA2325169-A1.
PN 03-JUN-2001.
XX
PF 01-DEC-2000; 2000CA-2325169.
XX
PR 03-DEC-1999; 99US-0168934.
XX
PA (NDSU-) NDSU RES FOUND.
PI Sheridan MA, Moore CA, Kittelson JD;
XX
DR WPI; 2001-425997/46.
DR N-PSDB; AAS12933.
XX
PT New somatostatin polypeptides derived from Oncoerhynchus mykiss, useful
PT for treating diabetes mellitus, acromegaly, gastrinoma, acquired
PT immunodeficiency syndrome and neurological disorders -
XX
PS Claim 1; Fig 2; 52pp; English.
XX
CC The invention relates to an Oncoerhynchus mykiss somatostatin polypeptide
CC containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of
CC preprosomatostatin II (PPSS-II). The protein sequences and their
CC associated polynucleotides are useful for identifying modified
CC somatostatin polypeptide which functions as a somatostatin agonist useful
CC for research, therapeutics or diagnostics, including medical and
CC veterinary applications. The wild-type somatostatin and its modified

CC version are useful for treating hypersecretion from endocrine tumours in
CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.
CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage
CC through their effects on cell proliferation and apoptosis and as adjuncts
CC in the treatment of diabetes mellitus via inhibition of growth hormone
CC and glucagon. In addition, dysfunctional somatostatin secretion is
CC associated with acquired immunodeficiency syndrome (AIDS) and various
CC neurological disorders (e.g. epilepsy, Alzheimer's disease and
CC Huntington's disease) and somatostatin antagonists are effective in the
CC treatment of such conditions. Nucleic acids encoding the polypeptides are
CC useful in gene therapy and fusion peptides can be targeted to neoplasms
CC and their metastases, inhibiting the release of their secretory products.
CC This sequence represents O. Mykiss PPSS-I protein.
CC Note: The features for this sequence are specifically claimed in the
CC specification.
XX
SQ Sequence 114 AA;
Query Match 27.6%; Score 164.5; DB 22; Length 114;
Best Local Similarity 38.7%; Pred. No. 8.5e-12;
Matches 48; Conservative 15; Mismatches 38; Indels 23; Gaps 5;
QY 1 MRVSQIHCAALLGLALAIICSGAASQPDLDLASRRLLQORALAAALPHRSGVSRWRTFY 60
Db 1 mlstrvqcalallslalaissvaaps---daklrqlqrslmap---agkqelarntl 53
QY 61 PNCPCLRWRPRKVKGPOLK-----AKEDLERSVDNLP-----PRRKAGCKNFWYWGK 107
Db 54 ve---llselahveneaeleddmshgveqevdelerapgvplaprerkagcknffwkt 110
QY 108 FTSC 111
Db 111 ftsc 114
RESULT 6
AAP61714
ID AAP61714 standard; Protein; 28 AA.
XX
AC AAP61714;
XX
DT 28-JUL-1991 (first entry)
XX
DE Somatostatin-28 analogue.
XX
KW Somatostatin-28; insulin-selective; insulinoma.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 23 /label= Hyl, Lys
FT
XX EP173527-A.
XX
PD 05-MAR-1986.
XX
PF 16-AUG-1985; 85EP-0305867.
XX
PR 31-AUG-1984; 84US-0646610.
XX
PA (SALK) SALK INST FOR BIOL STUD.
XX
PI Spiess J, Noe BD;
XX
DR WPI; 1986-063363/10.
XX
PT Angler fish somatostatin-28 and analogue and fragment - useful in
PT inhibiting insulin secretion in insulinoma.
XX
PS Claim 4; Page 18; 19pp; English.
XX

Qy	108	FTSC	111
Db	118	ftsc	121

RESULT 9

AA69789
ID AA69789 standard; Protein; 140 AA.

XX
AC AAY69789;

XX
DT 20-APR-2000 (first entry)

XX
DE MWPs-MWPmp20-(His)6-EGF-TEV-Somatostatin-28.

XX Fusion protein; Bacillus; cell wall protein; promoter; cleavage site;
KW TEV protease.

XX
QS Synthetic.

JP11341991-A.

XX
PD
14-DEC-1999.

XX
PF 30-MAR-1999; 99JP-0089488.

XX 31-MAR-1998: 98JP-0087339

31 MAR 1958, 300F 00087
 (ITOH-) ITOHAM FOODS INC
 PA

PA (UDAK//) UDAKA S.
XX

AA
PI Sato S, Higashikuni N, Kudo T, Kondo M;
XX

AX
DR WPI; 2000-101697/09.
DR N-PSDH. AA759209

OK N-PSDB; AAZC9209.

XX

PT A. DNA coding a new fused protein and preparation of a useful peptide through its expression -

PT

XX
PS Example 6; Fig 3; 43pp; Japanese.

The invention relates to a DNA construct encoding a fusion protein comprising a *Bacillus* species cell wall protein fused to a cleavage peptide and a heterologous protein. The fusion construct is placed downstream of a *Bacillus* species promoter sequence. This sequence represents the fusion protein MWpSP-MWPmp20-(His)6-EGR-TEV-somatostatin-28, an example of the fusion protein of the invention. It comprises the *Bacillus brevis* middle wall protein mp20 linked to the human somatostatin 28 protein via a (His)6-EGR linker-TEV protease cleavage site sequence.

XX	Sequence	140 AA;
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

Query Match	17.7%;	Score 105.5;	DB 21;	Length 140;
Best Local Similarity	28.0%;	Pred. No: 0.00012;		
Matches 42;	Conservative 18;	Mismatches 35;	Indels 55;	Gaps

QY **4** SQIHCA LALGLALA IC SGAA--SQPDLDASRRLLQRALAAALPHRSGVSEWRWTFYP 61
 | : |||| | : || : || : | : | : | : | : | : | :

Dz **4** svlasala itv apma faaeaaattt apkmdadmektvvhhh-----hhhsndse----- 51

Qy 62 NCP-----CL-----

22 0F10N9J0ZK9J0M}10U10N9J0E0N0V09J0
09 86 -B5VDNLP---PRERKAGCKNFYWKGETSC 111

ID AAP20198 standard; protein; 25 AA.

XX
AC

14-AUG-1992 (first entry)

14 AUG 1992 (1110Z GMT)
XX
DE Sequence of somatostatin-25 analogue.

sequence of somatostatin 25 analogs
XX
KW somatostatin: hormone: growth hormone

KM XX FH Rev

FT. Key
FT. Modified-sit
FT.

FI	Disulfide-bond	14..25	/table> n s
FT	Modified-site	25	

FT	Modified-site	25	/label= C-01
FT			
yy			

XX PN US4316891-A.

XX PD- 23-FEB-1982. YY

XX
PE 14-JUN-1980; 80US-0159801.
XX

XX
PR 14-JUN-1980; 80DS-0159801.

(SALK-) SALK INST BIOLOG.

Guillemin RCL, Esch FS, Bohlen P, Brazeau PE, Ling NC;
WPI; 1982-19801E/10.

Extended somatostatin analogues - with increased inhibition of growth hormone release

Claim 6; Column 12; 7pp; English.

The inventors claim a pharmaceutical compsn. which comprises (i) synthetic somatostatin-28 (SS-28), SS-25 or (D-Trp(22))-SS-28 and (ii) a liq. or solid carrier, and SS-28 derivs. and SS-25 derivs. The compsn. and derivs. are more potent then somatostatin in inhibiting release of growth hormone; they also inhibit basal and stimulated insulin and glucagon secretion. (D-Trp(22))-SS-28 exhibits very substantial increases in potency w.r.t. inhibition of growth hormone secretion.

XX	Sequence	25 AA;
SQ	Sequence	25 AA;

Query Match	17.6%
Best Local Similarity	81.0%

Best Local Matches

QY 91 LP PRKAGCKNFYWKGTSC 111

Db . . . 5 mar

11
PAGE 11

RESULT	11
AAB91017	
ID	AAB91017 standard; Peptide; 25 AA.
XX	
XX	
AC	AAB91017;
XX	
DT	22-JUN-2001 (first entry)
XX	
DE	Somatostatin related peptide SEQ ID NO:191.
XX	
KW	Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW	blood component; modification; succinimidyl; maleimido group; amino
KW	hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX
SO
XX

Homo sapiens.

Hydrolysis / emulsion

XX WO200069900-A2.
PN 23-NOV-2000.
XX 17-MAY-2000; 2000WO-US13576.
PF 17-MAY-1999; 99US-0134406.
XX 10-SEP-1999; 99US-0153406.
PR 15-OCT-1999; 99US-0159783.
XX (CONJ-) CONJUCHEM INC.
PA Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX WPI; 2001-112059/12.
XX
XX Modifying and attaching therapeutic peptides to albumin prevents
XX peptidase degradation, useful for increasing length of in vivo activity
XX
XX Disclosure; Page 252; 733pp; English.
XX
XX The present invention describes a modified therapeutic peptide (I)
XX comprising a therapeutically active amino acid region (III) and a
XX reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
XX a less therapeutically active amino acid region (IV), which covalently
XX bonds with amino/hydroxyl/thiol groups on blood components to form a
XX peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
XX (I) are useful for modifying therapeutic peptides e.g. hormones, growth
XX factors and neurotransmitters, to protect them from peptidase activity
XX in vivo for the treatment of various disorders. Endogenous therapeutic
XX peptides are not suitable as drug candidates as they require frequent
XX administration due to rapid degradation by peptidases in the body.
XX Modifying and attaching therapeutic peptides to albumin prevents or
XX reduces the action of peptidases to increase length of activity (half
XX life) and specificity as bonding to large molecules decreases
XX intracellular uptake and interference with physiological processes.
XX AAB90829 to AAB92441 represent peptides which can be used in the
XX exemplification of the present invention.
XX
XX Sequence 25 AA;
SQ

Query Match 17.6%; Score 105; DB 22; Length 25;
Best Local Similarity 81.0%; Pred. No. 1.7e-05;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Y 91 LPPRERKAGCKNFYWKGTSC 111
: |||||
db 5 maprerkgcknffwktftsc 25

RESULT 12
AAW51859
ID AAW51859 standard; peptide; 28 AA.
XX AAW51859;
XX
XX 10-SEP-1998 (first entry)
XX
XX Somatostatin analogue for the treatment of syndrome x of Reaven.
XX
XX Somatostatin analogue; syndrome x of Reaven; hyperinsulinaemia syndrome;
XX diazoxide; cyclothiazide; metformin.
XX Synthetic.
XX
XX WO9810786-A2.
XX
XX 19-MAR-1998.
XX
XX 10-SEP-1997; 97WO-IL00301.
PF

XX 10-OCT-1996; 96IL-0119403.
PR 12-SEP-1996; 96IL-0119250.
XX
XX (COHE/) COHEN Y.
XX Cohen Y;
XX
XX WPI; 1998-271636/24.
XX
XX Composition for treatment of the risk factors of syndrome x of
XX Reaven - (hyperinsulinaemia syndrome) comprises somatostatin,
XX diazoxide, cyclothiazide (or their analogues) and/or metformin
XX
XX Claim 42; Page 41; 45pp; English.
XX
XX The invention relates to a pharmaceutical composition for treatment of
XX the risk factors of syndrome x of Reaven (hyperinsulinaemia syndrome). It
XX comprises somatostatin, diazoxide, cyclothiazide (or an analogue of one
XX of these) or metformin as the active ingredient. The composition reduces
XX resistance to insulin, and so treats and prevents all the associated risk
XX factors at once. The risk factors are hypertension, dyslipidaemia
XX (raised triglyceride and LDL levels with reduced HDL levels), shorter
XX coagulation time due to increased Plasminogen Activator Inhibitor-1
XX levels, core obesity, glucose intolerance hyperinsulinaemia. The
XX composition reduces the incidence of ischaemic heart disease,
XX cerebrovascular disorders, intermittent claudication, ischaemic bowel
XX disease, impotence due to peripheral vascular disease, hypercoagulation
XX (e.g. renal vein thrombosis), obesity and glucose intolerance. The
XX present sequence represents a specifically claimed somatostatin analogue.
XX
XX Sequence 28 AA;
SQ

Query Match 17.6%; Score 105; DB 19; Length 28;
Best Local Similarity 81.0%; Pred. No. 2e-05;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 91 LPPRERKAGCKNFYWKGTSC 111
: |||||
Db 8 maprerkgcknffwktftsc 28

RESULT 13
AAY28703
ID AAY28703 standard; peptide; 28 AA.
XX
XX AAY28703;
XX
XX 07-OCT-1999 (first entry)
XX
XX Mouse somatostatin SS-28 hormone.

XX Mouse somatostatin SS-28 hormone; growth hormone; insulin; glucagon;
KW thyroid stimulating hormone; octreotide; cell-based delivery of insulin;
KW glucose-stimulated insulin secretion; SSTRV; somatostatin receptor;
KW mouse somatostatin receptor type V gene; diabetes.
XX Mus musculus.
XX
XX WO9935242-A1.
XX
XX 15-JUL-1999.
XX
XX 11-JAN-1999; 99WO-US00633.
XX
XX 03-JUN-1998; 98US-0087848.
PR 12-JAN-1998; 98US-0071193.
PR 12-JAN-1998; 98US-0071209.
PR 12-JAN-1998; 98US-0072556.
PR 03-JUN-1998; 98US-0087821.
XX
XX (BETA-) BETAGENE INC.
PA

XX Clark SA, Quaade C;
XX WPI; 1999-444195/37.
DR
XX New defined medium for culture of neuroendocrine cells, e.g. of
PT insulin-secreting cells
XX
XX Example 8; Page 143; 312pp; English.
PS
XX The present sequence is a mouse somatostatin (SS-28) hormone which
CC was found to inhibit the release of growth hormone, thyroid stimulating
CC hormone, insulin and glucagon. In addition, SS-28 and its analogue
CC Octreotide may inhibit growth of some tumours. The hormone was
CC used to study its effect on glucose-stimulated insulin secretion in
CC high expressing and non-expressing clones of mouse somatostatin
CC receptor, type V gene (SSRV). The insulin secretion was highly inhibited
CC in the high expressing clone as compared to the non-expressing clone
CC because the high expressing clone showed high sensitivity to
CC somatostatin. The hormone effectively inhibits insulin secretion in the
CC absence of glucose. The somatostatin receptor can be introduced in cell
CC lines, used in cell-based delivery of insulin for treating diabetes, for
CC precise regulation of insulin release.
XX
SQ Sequence 28 AA;

Query Match 17.6%; Score 105; DB 20; Length 28;
Best Local Similarity 81.0%; Pred. No. 2e-05;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 91 LPPRRKAGCKNFYWKGTSC 111
Db 8 maprrerkagcknffwktftsc 28

RESULT 14
AAY24384
ID AAY24384 standard; peptide; 28 AA.

XX AC AAY24384;
XX DT 20-SEP-1999 (first entry)
XX DE Somatostatin peptide hormone SS-28.
XX KW Glucagon-like peptide I receptor; GLP-1 receptor; drug screening;
XX secretory function; immortalized neuroendocrine secretory cell;
XX regulation; diabetes; insulin secretion; neuroendocrine-based disorder;
XX Parkinson's disease; athyrotic cretinism; Addison's disease.
XX OS Mus musculus.
XX PN WO9935495-A2.
XX PD 15-JUL-1999.
XX PF 11-JAN-1999; 99WO-US00551.
XX PR 03-JUN-1998; 98US-0087848.
XX PR 12-JAN-1998; 98US-0071193.
XX PR 12-JAN-1998; 98US-0071209.
XX PR 12-JAN-1998; 98US-0072556.
XX PR 03-JUN-1998; 98US-0087821.
XX (BETA-) BETAGENE INC.
XX PI Clark SA, Quaade C, Thigpen AE;
XX WPI; 1999-430454/36.
XX New modulators of secretory function, used to control peptide
PT secretion from cells in vivo or in vitro, specifically for treating

PT diabetes
XX Example 8; Page 153-154; 309pp; English.
PS
XX The present invention describes a method for identifying modulators (I)
CC of secretory function by treating an immortalized cell, having a stable
CC secretory function, with a test compound and detecting any change in
CC secretion caused by the compound. (I) are used to control secretion of
CC polypeptides from cells, in vivo or in vitro. Specifically they are used
CC for treating or preventing diabetes by regulation of insulin secretion,
CC but can also be used in cases of other neuroendocrine-based disorders
CC such as Parkinson's disease, athyrotic cretinism and Addison's disease.
CC The method uses engineered, immortalized cells that are available in
CC large amounts, with a stable and predictable phenotype. They allow
CC screening to be performed in vivo. The present sequence represents a
CC somatostatin peptide hormone SS-28 used in an example from the present
CC invention.
XX Sequence 28 AA;
SQ

Query Match 17.6%; Score 105; DB 20; Length 28;
Best Local Similarity 81.0%; Pred. No. 2e-05;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 91 LPPRRKAGCKNFYWKGTSC 111
Db 8 maprrerkagcknffwktftsc 28

RESULT 15
AAY24239
ID AAY24239 standard; peptide; 28 AA.

XX AC AAY24239;
XX DT 15-SEP-1999 (first entry)
XX DE Peptide hormone somatostatin SS-28.
XX KW Glucagon-like peptide I receptor; GLP-1; regulation; secretion;
XX neuroendocrine cell line; glycaemic sensing mechanism; glucose;
XX genetic engineering; hypoglycaemia; diabetes.
XX OS Synthetic.
XX PN WO9935255-A2.
XX PD 15-JUL-1999.
XX PF 11-JAN-1999; 99WO-US00631.
XX PR 03-JUN-1998; 98US-0087848.
XX PR 12-JAN-1998; 98US-0071193.
XX PR 12-JAN-1998; 98US-0071209.
XX PR 12-JAN-1998; 98US-0072556.
XX PR 03-JUN-1998; 98US-0087821.
XX (BETA-) BETAGENE INC.
XX PI Clark SA, Thigpen AE;
XX WPI; 1999-419351/35.
XX New immortalized neuroendocrine cells that stably secrete
PT polypeptide, particularly used to treat diabetes and hypoglycaemia
XX Example 8; Page 148; 318pp; English.
XX The present invention describes immortalized neuroendocrine cells (A)
CC that stably secrete a polypeptide hormone (I) contain an expression
CC region that includes a transgene (TG), linked to a promoter functional
CC in eukaryotic cells, such that expression of TG increases sensitivity

CC of the cells to a modulator of (I) secretion. (A) are specifically
CC used, by transplantation, to treat diabetes or hypoglycaemia (especially
CC where associated with insulin therapy) but more generally are used to
CC express, in vivo, a wide range of therapeutic hormones, enzymes,
CC amidated proteins and growth factors. Also engineered neuroendocrine
CC cells are used to identify new therapeutic agents or drug targets.
CC (A) have a stable phenotype and particularly inducible glucagon.
CC secretion and glucose counter-regulatory capacities, i.e. they balance
CC the hyperglycaemic effects of beta-cell loss and the hypoglycaemic
CC effects of administered insulin. Since they are of human origin, they
CC are less likely to suffer immune rejection than xenografts. The present
CC sequence represents the peptide hormone somatostatin SS-28 used in an
CC example from the present invention.

XX
SQ Sequence 28 AA;

Query Match 17.6%; Score 105; DB 20; Length 28;
Best Local Similarity 81.0%; Pred. No. 2e-05;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

91 LPPRRKAGCKNFYWKGTSC 111
:|||||:|||||
Db 8 maprerkagcknffwktftsc 28

Search completed: June 13, 2002, 12:22:21
Job time: 274 sec

QY 67 RWRPRKVGPKLAKKE-----DLERSVDNLP---PRERKAGCKNFWYKGFSTC 111
 Db 51 LSEPNOTENDALEPEDLPQAAEQDEMRLQRSANSNPAMAPRERKAGCKNFWKTFSTC 110

RESULT 2
 US-08-648-322-2
 ; Sequence 2, Application US/08648322
 ; Patent No. 6074872
 ; GENERAL INFORMATION:
 ; APPLICANT: Sutcliffe, Gregor J.
 ; APPLICANT: de Lecea, Luis
 ; TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
 ; STREET: 10666 No. 6074872th Torrey Pines Road, TPC-8
 ; CITY: La Jolla
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/648,322
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fitting, Thomas
 ; REGISTRATION NUMBER: 34,163
 ; REFERENCE/DOCKET NUMBER: 519.0
 ; TELEPHONE: (619) 554-2937
 ; TELEFAX: (619) 554-6312
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 112 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-648-322-2

Query Match 14.6%; Score 87; DB 3; Length 112;
 Best Local Similarity 25.2%; Pred. No. 0.0016;
 Matches 30; Conservative 9; Mismatches 42; Indels 38; Gaps 4;
 QY 20 CSQGAASQPDLDLASRRLLQRLAALPHRSGVSE-----WRTF 59
 Db 4 CSTRGKRPSALSLLLLLLSGIAASALPLESGPTGQDSVQDATGGRRTGLTFLAWH-- 61
 QY 60 YPNCPLRWRPRKVG-----POLKAKEDLERSVDNLPPEKAGCKNFWYKGFSTC 111
 Db 62 -----EWASQSSSTAFEGTPELSKQ--ERPPLOQPPHRDKKPKCKNFWKTFSSC 111

RESULT 3
 US-09-001-472-3
 ; Sequence 3, Application US/09001472
 ; Patent No. 6232100
 ; GENERAL INFORMATION:
 ; APPLICANT: OLSEN, HENRIK S.
 ; APPLICANT: RUBEN, STEVEN M.
 ; TITLE OF INVENTION: CORTISTATIN POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 ; STREET: 1100 NEW YORK AVENUE, SUITE 600

CITY: WASHINGTON
 STATE: DC
 COUNTRY: US
 ZIP: 20005-3934
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/001,472
 ; FILING DATE: Herewith
 ; CLASSIFICATION:
 ; PRIOR APPLICATION NUMBER: US 60/033,980
 ; FILING DATE: 31-DEC-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/037,386
 ; FILING DATE: 07-FEB-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: STEFFE, ERIC K.
 ; REGISTRATION NUMBER: 36,688
 ; REFERENCE/DOCKET NUMBER: 1488.0430002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 371-2600
 ; TELEFAX: (202) 371-2540
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 112 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-001-472-3

Query Match 14.6%; Score 87; DB 4; Length 112;
 Best Local Similarity 25.2%; Pred. No. 0.0016;
 Matches 30; Conservative 9; Mismatches 42; Indels 38; Gaps 4;
 QY 20 CSQGAASQPDLDLASRRLLQRLAALPHRSGVSE-----WRTF 59
 Db 4 CSTRGKRPSALSLLLLLLSGIAASALPLESGPTGQDSVQDATGGRRTGLTFLAWH-- 61
 QY 60 YPNCPLRWRPRKVG-----POLKAKEDLERSVDNLPPEKAGCKNFWYKGFSTC 111
 Db 62 -----EWASQSSSTAFEGTPELSKQ--ERPPLOQPPHRDKKPKCKNFWKTFSSC 111

RESULT 4
 US-08-648-322-6
 ; Sequence 6, Application US/08648322
 ; Patent No. 6074872
 ; GENERAL INFORMATION:
 ; APPLICANT: Sutcliffe, Gregor J.
 ; APPLICANT: de Lecea, Luis
 ; TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
 ; STREET: 10666 No. 6074872th Torrey Pines Road, TPC-8
 ; CITY: La Jolla
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/648,322
 ; FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: 519.0
TELEPHONE: (619) 554-2937
TELEFAX: (619) 554-6312
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: C-terminal
US-08-648-322-6

Query Match 13.9%; Score 83; DB 3; Length 85;
Best Local Similarity 44.4%; Pred. No. 0.0036;
Matches 16; Conservative 5; Mismatches 13; Indels 2; Gaps 1;
QY 76 POLKAKEDLERSVDNLPRLRKGCKNFYWKGTSC 111
Db 51 PELSKRQ--ERPPQQPPHRDRKPKCKNFFWKTSSC 84

RESULT 5
US-09-001-472-2
Sequence 2, Application US/09001472
Patent No. 6232100
GENERAL INFORMATION:
APPLICANT: OLSEN, HENRIK S.
APPLICANT: RUBEN, STEVEN M.
TITLE OF INVENTION: CORTISTATIN POLYPEPTIDES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: US
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/001,472
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,980
FILING DATE: 31-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/037,386
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0430002
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-001-472-2

Query Match 13.5%; Score 80.5; DB 4; Length 105;
Best Local Similarity 27.7%; Pred. No. 0.0098;
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;
QY 37 LLQRALA-AALPHRSGVSRWRTFYPNCPLR-----WRPRKVGKGPQL-KAK 81
Db 11 LLSGATAAALPLEGGPTGRDSEHMQEAGIRKSSLLTFLAWWFETWSAGPLIGEAA 70
QY 82 EDLERSVDNLP-----RERKAGCKNFYWKGTSC 111
Db 71 REVARROGAPPQQSARRDRMPCRNFFWKTSSC 104

RESULT 6
US-08-648-322-5
Sequence 5, Application US/08648322
Patent No. 6074872
GENERAL INFORMATION:
APPLICANT: Sutcliffe, Gregor J.
APPLICANT: de Lecea, Luis
TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
TITLE OF INVENTION: COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
STREET: 10666 NO. 6074872th Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/648,322
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: 519.0
TELEPHONE: (619) 554-2937
TELEFAX: (619) 554-6312
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-648-322-5

Query Match 13.3%; Score 79.5; DB 3; Length 109;
Best Local Similarity 31.2%; Pred. No. 0.014;
Matches 30; Conservative 4; Mismatches 35; Indels 27; Gaps 4;
QY 37 LLQRALAALPHRSG-----VSRWR-----TFYPNCPLRW-----RPRKVG 75
Db 19 LLGWAAASALPLESGTGDSDVQEAEGRSGLLTF-----LAWHHEWASQASSTPVGG 72
QY 76 POLKAKEDLERSVDNLPRLRKGCKNFYWKGTSC 111
Db 73 GTPGLSKSQERPPPPQQPPHLDKPKCKNFFWKTSSC 108

RESULT 7
US-08-648-322-7
Sequence 7, Application US/08648322
Patent No. 6074872

GENERAL INFORMATION:
APPLICANT: Sutcliffe, Gregor J.
APPLICANT: de Lecea, Luis
TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
TITLE OF INVENTION: COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
STREET: 10666 No. 6074872th Torrey Pines Road, TPC-8-
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/648,322
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: 519.0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 554-2937
TELEFAX: (619) 554-6312
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: C-terminal
US-08-648-322-7

Query Match 13.1%; Score 78; DB 3; Length 25
Best Local Similarity 51.9%; Pred. No. 0.0035;
Matches 14; Conservative 2; Mismatches 11; Indels

QY 85 ERSVDNLPERRKAGCKNFYWKGFESC 111
II II I IIIII I I I I
Db 2 ERPLQQPHRDKKPKCNFFWKTESSC 28

SULT 8
US-08-648-322-11
Sequence 11, Application US/08648322
Patent No. 6074872
GENERAL INFORMATION:
APPLICANT: Sutcliffe, Gregor J.
APPLICANT: de Lecea, Luis
TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
TITLE OF INVENTION: COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
STREET: 10666 No. 6074872th Torrey Pines Road, TPC-8-
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/648,322
FILING DATE:

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: 519.0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 554-2937
TELEFAX: (619) 554-6312
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: C-terminal
US-08-648-322-11

Query Match 13.1%; Score 7
Best Local Similarity 51.9%; Pred.N N
Matches 14; Conservative 2; Mismatches 0

QY 85 ERSVDNLPPrRKAGCKNFYWKGTSC 111
||| ||| ||| ||| ||| ||| ||| |||
DB 2 ERPPQPPhLDKKPKCNFFWKTSSC 28

RESULT 9
US-08-648-322-10
Sequence 10, Application US/08648322
Patent No: 6074872
GENERAL INFORMATION:
APPLICANT: Sutcliffe, Gregor J.
APPLICANT: de Lecea, Luis
TITLE OF INVENTION: CORTISTATIN: N
TITLE OF INVENTION: COMPOSITIONS AN
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: THE SCRIPPS RESEARCH I
STREET: 10666 NO. 6074872th Torre
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/648,322
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: 519.0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 554-2937
TELEFAX: (619) 554-6312
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: C-terminal
US-08-648-322-10

```

QY 85 ERSVDNLP--PRERKAGCKNFYWKFTSC 111
 Db 57 ERPPPPQPPHLDKPKCKNFYWKFTSSC 83

RESULT 10

US-08-455-970A-10
 ; Sequence 10, Application US/08455970A
 ; Patent No. 5708155
 ; GENERAL INFORMATION:
 ; APPLICANT: POTTER, ANDREW A.
 ; APPLICANT: REDMOND, MARK J.
 ; APPLICANT: HUGHES, HUW P.A.
 ; TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: REED & ROBINS
 ; STREET: 285 HAMILTON AVENUE, SUITE 200
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 94301

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/455,970A
 FILING DATE: 31-MAY-1995
 CLASSIFICATION: 424

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/960,932
 FILING DATE: 14-OCT-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: ROBINS, ROBERTA L.
 REGISTRATION NUMBER: 33,208
 REFERENCE/DOCKET NUMBER: 9001-0016.10
 TELEPHONE: (415) 327-3400
 TELEFAX: (415) 327-3231
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 943 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-455-970A-10

Query Match 13.0%; Score 77.5; DB 1; Length 943;
 Best Local Similarity 48.4%; Pred. No. 0.46;
 Matches 15; Conservative 5; Mismatches 8; Indels 3; Gaps 1;

QY 84 LERSVDNLP---PRERKAGCKNFYWKFTSC 111
 Db 913 LDQSLSSLFQFARGSSSSAGCKNFYWKFTTSC 943

RESULT 11

US-07-977-628A-1
 ; Sequence 1, Application US/07977628A
 ; Patent No. 5405597
 ; GENERAL INFORMATION:
 ; APPLICANT: Dean, Richard T
 ; APPLICANT: Lister-James, John
 ; APPLICANT: Buttram, Scott
 ; TITLE OF INVENTION: Technetium-99m Labeled Somatostatin-
 ; DERIVED PEPTIDES FOR IMAGING
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.
 STREET: 10 South Wacker Drive, Suite 3000
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/977,628A
 FILING DATE: 17-NOV-1992
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5405597nan, Kevin E
 REGISTRATION NUMBER: 35,303
 REFERENCE/DOCKET NUMBER: 91,642-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-715-1000
 TELEFAX: 312-715-1234
 TELEX: 910-221-5317
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 14 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 2..14
 OTHER INFORMATION: /label= Disulfide-bond
 OTHER INFORMATION: /note= "The sidechain thiol groups of the third
 OTHER INFORMATION: residue cysteine and the carboxy-terminal cysteine
 OTHER INFORMATION: form a disulfide bond in native somatostatin
 US-07-977-628A-1

Query Match 12.9%; Score 77; DB 1; Length 14;
 Best Local Similarity 85.7%; Pred. No. 0.0018;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYWKFTSC 111
 Db 1 AGCKNFYWKFTSC 14

RESULT 12

US-08-255-272-5
 ; Sequence 5, Application US/08255272
 ; Patent No. 5627268
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Ramesh
 ; APPLICANT: Sharma, Ajay
 ; APPLICANT: Khoury-Christianson, Anastasia
 ; APPLICANT: M.
 ; TITLE OF INVENTION: Production of Therapeutic Peptides in
 ; TRANSGENIC ANIMALS AS A FUSION WITH HEMOGLOBIN
 ; NUMBER OF SEQUENCES: 32
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PENNIE & EDMONDS
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/255,272
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Coruzzi, Laura A.
 REGISTRATION NUMBER: 30742
 REFERENCE/DOCKET NUMBER: 6794-032
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 14 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-255-272-5

Query Match 12.9%; Score 77; DB 1; Length 14;
 Best Local Similarity 85.7%; Pred. No. 0.0018;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYWKGTSC 111
 Db 1 AGCKNFFWKFTSC 14

RESULT 13
 US-08-416-007-4
 Sequence 4, Application US/08416007
 Patent No. 5693679
 GENERAL INFORMATION:
 APPLICANT: Vincent, Jean-Pierre
 APPLICANT: Gaudriault, Georges
 APPLICANT: Beaudet, Alain
 TITLE OF INVENTION: FLUORESCENT SOMATOSTATIN
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/416,007
 FILING DATE: 04-APR-1995
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 06942/003001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 14 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-416-007-4

Query Match 12.9%; Score 77; DB 1; Length 14;
 Best Local Similarity 85.7%; Pred. No. 0.0018;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYWKGTSC 111
 Db 1 AGCKNFFWKFTSC 14

RESULT 14
 US-08-676-263-11
 Sequence 11, Application US/08676263
 Patent No. 5705143
 GENERAL INFORMATION:
 APPLICANT: Bower, Gary R.
 APPLICANT: Forster, Alan M.
 APPLICANT: Riley, Anthony L. M.
 APPLICANT: Storey, Anthony E.
 TITLE OF INVENTION: BIOLOGICAL TARGETING AGENTS
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 233 South Wacker Drive/6300 Sears Tower
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/676,263
 FILING DATE: 07-NOV-1996
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: EP 94300224.6
 FILING DATE: 12-JAN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Sharp, Jeffrey S.
 REGISTRATION NUMBER: 31,879
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX: (312) 474-6600
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 14 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant
 MOLECULE TYPE: peptide
 FEATURE:
 NAME/KEY: Cross-links
 LOCATION: 3..14
 US-08-676-263-11

Query Match 12.9%; Score 77; DB 1; Length 14;
 Best Local Similarity 85.7%; Pred. No. 0.0018;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYWKGTSC 111
 Db 1 AGCKNFFWKFTSC 14

RESULT 15
 US-08-286-748B-13
 Sequence 13, Application US/08286748B
 Patent No. 5759542

GENERAL INFORMATION:
APPLICANT: Victor Gurewich
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
OF DRUGS BY PLATELETS FOR THE TREATMENT OF
CARDIOVASCULAR AND OTHER DISEASES
TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
CARDIOVASCULAR AND OTHER DISEASES
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,748B
FILING DATE: August 5, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: J. Peter Fasse
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04547/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 14
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-286-748B-13

Query Match 12.9%; Score 77; DB 1; Length 14;
Best Local Similarity 85.7%; Pred. No. 0.0018;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 98 AGCKNFWKGTSC 111
|||||:|||||
1 AGCKNFWKGTSC 14

Search completed: June 13, 2002, 12:22:56
Job time: 194 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using .sw model

Run on: June 13, 2002, 12:29:53 ; Search time 18.73 Seconds
(without alignments) 229.464 Million cell updates/sec

Title: US-09-727-739B-15
Perfect score: 597
Sequence: 1 MRVSIHCALALGLALAIC.....PPRRKAGCKNFYWKGTSC 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 1052224

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 s
```

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	476	79.7	115	1	SMS2_ONCMY	Q91194 oncorhynchu
2	205.5	34.4	120	1	SMS2_CARAU	Q9Ygh4 carassius a
3	189	31.7	125	1	SMS2_IOPAM	P01170 lophius ame
4	179.5	30.1	116	1	SMS_CHICK	P33094 gallus gall
5	178.5	29.9	114	1	SMSA_CARAU	Q9Ygh5 carassius a
6	176.5	29.6	116	1	SMS_HUMAN	P01166 homo sapien
7	174	29.1	115	1	SMS1_RANRI	P87384 rana ridibu
8	166.5	27.9	116	1	SMS_SHEEP	Q46688 ovis aries
9	163.5	27.4	116	1	SMS_MOUSE	P01167 mus musculu
10	162.5	27.2	116	1	SMS_BOVIN	P26917 bos taurus
11	160.5	26.9	116	1	SMS_CANFA	P49670 canis fami
12	156.5	26.2	114	1	SMS1_ICTPU	P01171 ictalurus p
13	153	25.6	115	1	SMS1_PROAN	Q9W7f0 protopteru
14	136	22.8	73	1	SMS2_PLAFE	P21780 platichthy
15	136	22.8	92	1	SMS_PIG	P01168 sus scrofa
16	131.5	22.0	121	1	SMS1_LOPAM	P01169 lophius ame
17	125.5	21.0	74	1	SMS2_MYOSC	P09876 myoxocephal
18	121	20.3	28	1	SMS2_ORENI	P81029 oreochromis
19	107	17.9	26	1	SMS1_AMICA	Q9prz6 amia calva
20	104	17.4	34	1	SMS_MYXGL	P19209 myxine glut
21	102.5	17.2	111	1	SMSB_CARAU	Q9Ygh3 carassius a
22	100.5	16.8	109	1	SMS2_PROAN	Q9W7e9 protopteru
23	94	15.7	37	1	SMS_PETWA	P21779 petromyzon
24	93	15.6	35	1	SMS_LANFL	Q9prro lampetra fl
25	91.5	15.3	103	1	SMS2_RANRI	P87385 rana ridibu
26	87	14.6	112	1	CORT_RAT	Q62949 rattus norv
27	80.5	13.5	105	1	CORT_HUMAN	Q00230 homo sapien
28	79.5	13.3	109	1	CORT_MOUSE	P56469 mus musculu
29	77.5	13.0	1236	1	POLS_WEEV	P13897 western equ
30	77	12.9	14	1	SMS1_MYOSC	P20750 myoxocephal
31	77	12.9	14	1	SMS_ALLMI	P31885 alligator m
32	74.5	12.5	234	1	YARE_CORGL	P38118 corynebacte
33	73.5	12.3	342	1	CYS2_HAEEO	P25793 haemonechus

	70	11.7	314	1	Y4RB_RHISN	P55635	rhizobium s
34	69.5	11.6	342	1	CYS1_HAECO	P19092	haemochus
35	69.5	11.6	1305	1	GAK_RAT	P97874	rattus norv
36	69.5	11.6	1204	1	NOS3_BOVIN	P29473	bos taurus
37	69	11.6	399	1	BM8A_MOUSE	P34821	mus musculus
38	68.5	11.5	133	1	SY21_MOUSE	O09006	mus musculus
39	67	11.2	260	1	PT22_SACBA	O13374	saccharomyc
40	67	11.2	1204	1	NOS3_PIG	Q28969	sus scrofa
41	66	11.1	317	1	CAH6_MOUSE	P18761	mus musculus
42	64.5	10.8	105	1	SMS2 ICTPU	P01172	ictalurus p
43	64	10.7	988	1	E4L2_MOUSE	O70318	mus musculus
44	64	10.7	100	1	APA2_MACFA	P18656	macaca fasc
45	62.5	10.5,					

ALIGNMENTS

[illegible]

```
Db 61 PNCPL--RPRKVKCP-AGAKEDLRVELERSVGNPNLPPRRKAGCKNFYWKGTSC 115

RESULT 2
SMS2_CARAU
ID SMS2_CARAU STANDARD; PRT; 120 AA.
AC Q9YGH4; Q9PTU2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28;
[Tyr7,Gly10]somatostatin-14].
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Ostariophysi;
OC Cypriniformes; Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
TISSUE=Brain;
Lin X.-W., Peter R.E.;
"Cloning and characterization of cDNAs encoding preprosomatostatin-I
and -II from goldfish brain.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
TISSUE=Liver;
Otto C.J., Peter R.E.;
"The expression of SRIF mRNA in the brain of goldfish.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U60262; AAD09626.1;
DR EMBL; AF025686; AAF15306.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 92 POTENTIAL.
FT PEPTIDE 93 120 [TYR21, GLY24]SOMATOSTATIN-28.
FT DISULFID 107 120 [TYR7, GLY10]SOMATOSTATIN-14.
FT CONFLICT 109 120 BY SIMILARITY.
FT CONFLICT 51 51 Q -> RW (IN REF. 2).
SQ SEQUENCE 120 AA; 13723 MW; 98957D68011A651A CRC64;

Query Match 34.4%; Score 205.5; DB 1; Length 120;
Best Local Similarity 42.0%; Pred. No. 2.6e-15;
Matches 55; Conservative 14; Mismatches 31; Indels 31; Gaps 7;

QY 1 MRYSQIHCALALLGLALAICSGAASQ--PDLDLARRLLQRLAALPHRSGVSRWRT 58
||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MRCELHCYLALLGLSLVLCGRCAANSQLEPDLFRHRLIQA-----SATQATQD 52

QY 59 FYP-----NCPCLRRPRKVKGPQLKAK-EDL-----ERSVDN---LPPRRKAGC 100
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 53 FTKRDVEKLLSLISIPMEMRE---KGLSMAGESEDLRLQEERSAESNLPTVRKEGC 109

QY 101 KNEYWKGTSC 111
|||||
Db 110 KNEYWKGTSC 120
|||||
```

```
RESULT 3
SMS2_LOPAM
ID SMS2_LOPAM STANDARD; PRT; 125 AA.
AC P01170; Q91066;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr7,Gly10]somatostatin-14].
OS Lophius americanus (American goosefish) (Anglerfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Lophiiformes; Lophidae; Lophius.
OX NCBI_TaxID=8073;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE-81052423; PubMed-6107860;
RA Hobart P.M., Crawford R., Shen L., Pictet R., Rutter W.J.;
"Cloning and sequence analysis of cDNAs encoding two distinct
somatostatin precursors found in the endocrine pancreas of
anglerfish.";
RL Nature 288:137-141(1980).
[2]
RP PARTIAL SEQUENCE, AND HYDROXYLATION.
RX MEDLINE-87308304; PubMed-2887572;
RA Andrews P.C., Nichols R., Dixon J.E.;
"Post-translational processing of preprosomatostatin-II examined
using fast atom bombardment mass spectrometry.";
RL J. Biol. Chem. 262:12692-12699(1987).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: SOMATOSTATIN II MAY HAVE A DIFFERENT DEGREE OF
CC ACTIVITY OR A DIFFERENT TYPE OF TARGET CELL FROM SOMATOSTATIN I.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; V00641; CAA23987.1;
DR PIR; A01434; RIAFS2.
DR PIR; A27376; A27376.
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal; Hydroxylation;
KW Multigene family.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 109 [TYR7, GLY10]SOMATOSTATIN-14.
FT PEPTIDE 112 125
FT DISULFID 114 125
FT MOD_RES 120 120 HYDROXYLATION.
FT CONFLICT 77 78 DV -> TG (IN REF. 1).
FT CONFLICT 90 90 G -> E (IN REF. 1).
SQ SEQUENCE 125 AA; 14052 MW; 5E14605D7B9A46FE CRC64;

Query Match 31.7%; Score 189; DB 1; Length 125;
Best Local Similarity 39.7%; Pred. No. 1.6e-13;
Matches 52; Conservative 13; Mismatches 32; Indels 34; Gaps 7;

QY 6 IHCALALLGLALAICSGAASQ-----PDLDLARR--LQRLAALPHRSGVSR 55
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4 IRCPAILALLVLCGSPSVSSQLDREQSDNQDLRLQHWLLERARSAGL-----LSQE 58

QY 56 W-----RTFFPNCPLRRPRKVKGPQLK--AKEDLRSVD---NLPPRRKAGC 100
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 59 WSKRAVEELLAQMSLPADV-----QREAEASMATGGRMNLRSVDSTNNLPPRRKAGC 114

QY 101 KNEYWKGTSC 111
|||||
```


RT "Sequence of the human somatostatin I gene.";
 RL Science 224:168-171(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RX MEDLINE=83014931; PubMed=6126875;
 RA Shen L.-P., Pictet R.L., Rutter W.J.;
 RT "Human somatostatin I: sequence of the cDNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:4575-4579(1982).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.fascicularis;
 RX MEDLINE=88144503; PubMed=2894033;
 RA Travis G.H., Sutcliffe J.G.;
 RT "Phenol emulsion-enhanced DNA-driven subtractive cDNA cloning:
 RT isolation of low-abundance monkey cortex-specific mRNAs.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:1696-1700(1988).
 CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PHARMACEUTICAL: Available under the name Sandostatin (Novartis);
 CC this is a synthetic cyclic analog known as octreotide or SMS
 CC 201-995. Used for the treatment of a variety of disorders
 CC including acromegaly and the symptomatic treatment of carcinoid
 CC tumors and vasoactive intestinal peptide tumors.
 CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J00306; AAA60566.1; -
 DR EMBL; M19318; AAA36908.1; -
 DR PIR; A33614; RIHUS1.
 DR PIR; A28968; A28968.
 DR MIM; 182450; -
 DR InterPro; IPR004250; Somatostatin.
 DR Pfam; PF03002; Somatostatin; 1.
 KW Cleavage on pair of basic residues; Hormone; Signal; Pharmaceutical.
 FT SIGNAL 1 24
 FT PROPEP 25 88
 FT PEPTIDE 89 116 SOMATOSTATIN-28.
 FT PEPTIDE 103 116 SOMATOSTATIN-14.
 FT DISULFID 105 116
 FT SEQUENCE 116 AA; 12735 MW; AB49BB89DC9DD8DA CRC64;
 SQ
 Query Match 29.6%; Score 176.5; DB 1; Length 116;
 Best Local Similarity 38.2%; Pred. No. 3.3e-12;
 Matches 47; Conservative 19; Mismatches 30; Indels 27; Gaps 6;
 QY 5 QIHCALALLGLALAI-CSQGAASQPDLDLASRRLLQRALAAALPHRSGVSEWRWTFYFNC 63
 Db 5 RLQCALAALSIVLALGCVTGAPSDPRL----RQFLQKSLAAA---AGKQELAKYFLAE- 55
 QY 64 PCLWRPRKVKGPQLKAKE-----DLERSVDNLP---PRERKAGCKNFYWKGF 108
 Db 56 --LLSEPNQTNDALEPEDLSQAABQDEMRLQRSANSNPAMPAPRERKAGCKNFYWKTF 113
 QY 109 TSC 111
 Db 114 TSC 116

RESULT 7
 SMS1_RANRI
 ID SMS1_RANRI STANDARD; PRT; 115 AA.
 AC P87384; Q9PSI8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Somatostatin I precursor (PSS1) [Contains: Somatostatin-14 (S-I)
 DE (SSSI)].
 OS Rana ridibunda (Laughing frog) (Marsh frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.
 OX NCBI_TaxID=8406;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97057290; PubMed=8901629;
 RA Tostivint H., Lihmann I., Bucharies C., Vleau D., Coulouarn Y.,
 RA Fournier A., Conlon J.M., Vaudry H.;
 RT "Occurrence of two somatostatin variants in the frog brain:
 RT characterization of the cDNAs, distribution of the mRNAs, and
 RT receptor-binding affinities of the peptides.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:12605-12610(1996).
 RN [2]
 RP SEQUENCE OF 102-115.
 RC TISSUE=Brain;
 RX MEDLINE=93038702; PubMed=1358069;
 RA Vaudry H., Chartrel N., Conlon J.M.;
 RT "Isolation of [pro2,Met13]somatostatin-14 and somatostatin-14 from the
 RT frog brain reveals the existence of a somatostatin gene family in a
 RT tetrapod.";
 RL Biochem. Biophys. Res. Commun. 188:477-482(1992).
 CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U68136; AAC60093.1; -
 DR HSSP; P02619; 3PAL.
 DR InterPro; IPR004250; Somatostatin.
 DR Pfam; PF03002; Somatostatin; 1.
 KW Cleavage on pair of basic residues; Hormone; Multigene family; Signal.
 FT SIGNAL 1 24
 FT PROPEP 25 99
 FT PEPTIDE 102 115 BY SIMILARITY.
 FT DISULFID 104 115 SOMATOSTATIN-14.
 FT SEQUENCE 115 AA; 12691 MW; 349756FEB4ABE213 CRC64;
 SQ
 Query Match 29.1%; Score 174; DB 1; Length 115;
 Best Local Similarity 39.4%; Pred. No. 6.1e-12;
 Matches 50; Conservative 17; Mismatches 32; Indels 28; Gaps 6;
 QY 1 MRVSIHICALALLGLALAI-CSQGAASQPDLDLASRRLLQRALAAALPHRSGVSEWRWTF 59
 Db 1 MQSCRVQCALTLLSLALAINISAAPTDPRL----RQFLQKSLASA-----GKQELAKYF 51
 QY 60 YPNCPCLRWRPRKVKGPQLKAKE-----DLERSVDNLP---PRERKAGCKNFY 104
 Db 52 LAE---LLSEPSQTDNEALESDDLPRGAEQDEVRLERSANSNPALAPRERKAGCKNF 108
 QY 105 WKFTSC 111
 Db 109 WKFTSC 115

RESULT 8
 SMS_SHEEP
 ID SMS_SHEEP STANDARD; PRT; 116 AA.
 AC O46688;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)

DR EMBL; K02248; AAA42161.1; -
DR EMBL; V01271; CAA24579.1; -
DR EMBL; J00787; AAA42164.1; -
DR EMBL; M25890; AAA42167.1; -
DR EMBL; J00788; AAA42162.1; -
DR EMBL; X51468; CAA35831.1; -
DR PIR; A20983; RIRTS1.
DR PIR; S08416; S08416.
DR MGD; MGI:98326; Smst.
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal.
FT SIGNAL 1 24 ANTRIN.
FT PEPTIDE 25 34
FT PROPEP 35 88
FT PEPTIDE 89 116 SOMATOSTATIN-28.
FT PEPTIDE 103 116 SOMATOSTATIN-14.
FT DISULFID 105 116
FT CONFLICT 43 43 T -> Y (IN REF. 5).
FT CONFLICT 79 79 Q -> H (IN REF. 6).
-T SEQUENCE 116 AA; 12745 MW; D48B5454C4490375 CRC64;

Query Match 27.4%; Score 163.5; DB 1; Length 116;
Best Local Similarity 37.4%; Pred. No. 8.3e-11;
Matches 46; Conservative 19; Mismatches 31; Indels 27; Gaps 6;

QY 5 QIHCALALLGLALAICS-QGAASQPDLDLASRLLQRLAALPHRSGVSRWRTFYPNC 63
Db 5 RLQCALAALCIIVLALGGVTGAPSDPRL-----RQFLQKSLAA-----TGKQELAKYFLAE- 55

QY 64 PCLWRPRKVKGPQLKAKE-----DLERSVDNLP-----PRERKAGCKNFYWKGF 108
Db 56 --LLSEPNQTEINDALEPEDLPQAAEQDEMRLQRSANSNPANAPRERKAGCKNFYWKTF 113

QY 109 TSC 111
Db 114 TSC 116

RESULT 10
SMS_BOVIN
ID SMS_BOVIN STANDARD; PRT; 116 AA.
AC P26917;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].
GN SST.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88288237; PubMed=2899837;
RA Su C.J., White J.W., Li W.H., Luo C.C., Frazier M.L., Saunders G.F.,
Chap L.;
RT "Structure and evolution of somatostatin genes.";
RL Mol. Endocrinol. 2:209-216(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-HOLSTEIN;
RX MEDLINE=99198780; PubMed=10100681;
RA Furu L.M., Kazmer G.W., Strausbaugh L., Zinn S.A.;
RT "Cloning and characterization of the bovine somatostatin gene.";
RL J. Anim. Sci. 77:492-493(1999).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M31217; AAA30744.1; -
DR EMBL; U97077; AAB58056.1; -
DR PIR; A40929; A40929.
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal.
FT SIGNAL 1 24 BY SIMILARITY.
FT PROPEP 25 88 BY SIMILARITY.
FT PEPTIDE 89 116 SOMATOSTATIN-28.
FT PEPTIDE 103 116 SOMATOSTATIN-14.
FT DISULFID 105 116 BY SIMILARITY.
SQ SEQUENCE 116 AA; 12688 MW; C18F17E64A371D8E CRC64;

Query Match 27.2%; Score 162.5; DB 1; Length 116;
Best Local Similarity 37.4%; Pred. No. 1.1e-10;
Matches 46; Conservative 19; Mismatches 31; Indels 27; Gaps 6;

QY 5 QIHCALALLGLALAICS-QGAASQPDLDLASRLLQRLAALPHRSGVSRWRTFYPNC 63
Db 5 RLQCALAALSIIVLALGGVTGAPSDPRL-----RQFLQKSLAA-----AGKQELAKYFLAE- 55

QY 64 PCLWRPRKVKGPQLKAKE-----DLERSVDNLP-----PRERKAGCKNFYWKGF 108
Db 56 --LLSEPNQTEINDALEPEDLPQAAEQDEMRLQRSANSNPANAPRERKAGCKNFYWKTF 113

QY 109 TSC 111
Db 114 TSC 116

RESULT 11
SMS_CANFA
ID SMS_CANFA STANDARD; PRT; 116 AA.
AC P49670;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].
GN SST.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gastric mucosa;
RX MEDLINE=97142297; PubMed=8988514;
RA Dickinson C.J., Delvalle J., Todisco A., Gantz I., Tong L.,
Finniss S., Yamada T.;
RT "Canine prosomatostatin: isolation of a cDNA, regulation of gene
expression; and characterization of post-translational processing
intermediates.";
RL Regul. Pept. 67:145-152(1996).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL; L42325; AAA67099.1; -
 DR InterPro; IPR004250; Somatostatin.
 DR Pfam; PF03002; Somatostatin; 1.
 KW Cleavage on pair of basic residues; Hormone; Signal.
 FT SIGNAL 1 24 BY SIMILARITY.
 FT PROPEP 25 88 BY SIMILARITY.
 FT PEPTIDE 89 116 SOMATOSTATIN-28.
 FT PEPTIDE 103 116 SOMATOSTATIN-14.
 FT DISULFID 105 116 BY SIMILARITY.
 SQ SEQUENCE 116 AA; 12735 MW; AB49BD231E731C9E CRC64;

Query Match 26.9%; Score 160.5; DB 1; Length 116;
 Best Local Similarity 36.5%; Pred. No. 1.8e-10;
 Matches 46; Conservative 19; Mismatches 28; Indels 33; Gaps 7;

QY 5 QIHCALALLGLALAI---CSQGAASQPDLDLASRRLLQALAAALPHRSGVSRWTFY 60
 Db 5 RLQCALAALSIVLALGGVTC---APSDPRL---RQFLQKSLAA---AGKQELAKYFL 53
 Y 61 PNCPLRWPRKVKGPOLKAKE-----DLERSVDNLP---PRKAGCKNFYW 105
 Db 54 AE---LLSEPQNTENDALEPEDLSQAAEQDEMRLQRSANSNPAMAPRERKAGCKNFFW 110
 QY 106 KGFTSC 111
 Db 111 KTFTSC 116

RESULT 12
 SMS1 ICTPU STANDARD; PRT; 114 AA.
 AC P01171;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Somatostatin I precursor [Contains: Somatostatin-14 (SS-14)].
 OS Ictalurus punctatus (Channel catfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Siluriformes; Ictaluridae; Ictalurus.
 OX NCBI_TaxID=7998;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85303576; PubMed=2863931;
 RA Dixon J.E., Andrews P.C.;
 RT "Somatostatins of the channel catfish."
 L Adv. Exp. Med. Biol. 188:19-29(1985).
 N [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=82265698; PubMed=6179939;
 RA Minth C.D., Taylor W.L., Magazin M.D., Taviani M.A., Collier K.J.,
 RA Weith H.L., Dixon J.E.;
 RT "The structure of cloned DNA complementary to catfish pancreatic
 somatostatin-14 messenger RNA."
 RL J. Biol. Chem. 257:10372-10377(1982).
 RN [3]
 RP SEQUENCE OF 82-114 FROM N.A.
 RX MEDLINE=82082515; PubMed=6171821;
 RA Taylor W.L., Collier K.J., Deschenes R.J., Weith H.L., Dixon J.E.;
 RT "Sequence analysis of a cDNA coding for a pancreatic precursor to
 somatostatin."
 RL Proc. Natl. Acad. Sci. U.S.A. 78:6694-6698(1981).
 RN [4]
 RP SEQUENCE OF 101-114.
 RX MEDLINE=81264223; PubMed=6114953;
 RA Andrews P.C., Dixon J.E.;
 RT "Isolation and structure of a peptide hormone predicted from a mRNA
 sequence. A second somatostatin from the catfish pancreas."
 RL J. Biol. Chem. 256:8267-8270(1981).
 CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: PANCREAS.

CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M25903; AAA49339.1; -
 DR EMBL; V00607; CAA23877.1; -
 DR EMBL; V00608; CAA23878.1; -
 DR PIR; S00292; RIIDS1.
 DR InterPro; IPR004250; Somatostatin.
 DR Pfam; PF03002; Somatostatin; 1.
 KW Cleavage on pair of basic residues; Hormone; Signal;
 KW Multigene family.
 FT SIGNAL 1 24 PROBABLE.
 FT PEPTIDE 101 114 SOMATOSTATIN-14.
 FT DISULFID 103 114
 FT CONFLICT 62 62
 SQ SEQUENCE 114 AA; 12419 MW; FEE0F2C76F74D99F CRC64;

Query Match 26.2%; Score 156.5; DB 1; Length 114;
 Best Local Similarity 31.4%; Pred. No. 4.6e-10;
 Matches 43; Conservative 20; Mismatches 25; Indels 49; Gaps 5;

QY 1 MRYSQIHCALALLGLALAICS-QGAASQPDLDLASRRLLQALAAALPHRSGVSRWTF 59
 Db 1 MPSTRIQCALALLAVALSVCSVSGAPS---DAKLQFLQRSILA----- 41
 QY 60 YPNCPLRWPRKVKGPOLKAK-----EDLERSVDN-----LPPR 94
 Db 42 ----PSVKQELTRYTLAELLAELAENEVLDSDEVSRAAESEGARLEMERAAAGPMAPR 97
 QY 95 ERKAGCKNFYWKGTSC 111
 Db 98 ERKAGCKNFFWKFTSC 114

RESULT 13
 SMS1 PROAN STANDARD; PRT; 115 AA.
 AC Q9W7F0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Somatostatin I precursor (PSSI) [Contains: Somatostatin-27;
 DE Somatostatin-14].
 OS Protopterus annectens (African lungfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Dipnoi; Lepidosireniiformes; Protopteridae; Protopterus.
 OX NCBI_TaxID=7888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93326690; PubMed=10398054;
 RA Trabucchi M., Tostivint H., Lihmann I., Jegou S., Vallarino M.,
 RA Vaudry H.;
 RT "Molecular cloning of the cDNAs and distribution of the mRNAs encoding
 two somatostatin precursors in the African lungfish Protopterus
 annectens."
 RL J. Comp. Neurol. 410:643-652(1999).
 CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way


```
DR PIR; A01432; RIGPS.
DR PIR; A24222; A24222.
DR PIR; A34109; A34109.
DR PIR; S13616; S13616.
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone.
FT NON_TER 1 1
FT PROPEP 1 64
FT PEPTIDE 65 92 SOMATOSTATIN-28.
FT PEPTIDE 79 92 SOMATOSTATIN-14.
FT DISULFID 81 92
SQ SEQUENCE 92 AA; 10346 MW; 787CBE82CFBBAE76 CRC64;

Query Match 22.8%; Score 136; DB 1; Length 92;
Best Local Similarity 36.9%; Pred. No. 5.9e-08;
Matches 38; Conservative 15; Mismatches 24; Indels 26; Gaps 5;

y 24 AASQPDLDLASRRLLQRLAAALPHRSVSEWRWTFYPCPCLRWRPRKVKGPOLKAKE- 82
Db 1 APSDPRL----RQFLQKSLAAA---AGKQELAKYFLAE---LLSEPNQTNDALEPEDL 49

QY 83 -----DLERSVDNLP---PRERKAGCKNFYWKGTSC.111
Db 50 SQAEEQDENRLELQRSANSNPAMAPRERKAGCKNFFWKTFITSC 92
```

Search completed: June 13, 2002, 12:29:54
Job time: 446 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:31:02 ; Search time 60.72 Seconds
(without alignments)
31.340 Million cell updates/sec

Title: US-09-727-739B-18
Perfect score: 58
Sequence: 1 SVDNLPPEPRK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	79.3	115	13	Q90Y43 osteoglossu
2	46	79.3	845	5	Q9V466 drosophila
3	42	72.4	273	3	O13787 schizosacch
4	41	70.7	114	13	Q90Y42 pantodon bu
5	40	69.0	242	4	Q9V3T1 homo sapien
6	40	69.0	275	16	Q98BW0 rhizobium.1
7	40	69.0	349	16	Q92WH3 rhizobium m
8	40	69.0	368	17	Q97YY4 sulfolobus
9	40	69.0	477	2	Q938D1 mycobacteri
10	40	69.0	962	5	Q9N6U5 drosophila
11	40	69.0	1006	5	Q9VTE2 drosophila
12	39	67.2	219	10	Q9FQE4 glycine max
13	39	67.2	224	10	Q9FQD8 glycine max
14	39	67.2	444	16	Q9RX02 deinococcus
15	39	67.2	796	5	Q9VAP3 drosophila
16	38	65.5	158	11	Q9CYE7 mus musculu

17	38	65.5	175	4	Q9HAA3	Q9haa3 homo sapien
18	38	65.5	236	10	Q9FYB0	Q9fyb0 arabidopsis
19	38	65.5	367	17	Q972M6	Q972m6 sulfolobus
20	38	65.5	486	5	Q9VQT3	Q9vqt3 drosophila
21	38	65.5	490	5	Q9V7Y5	Q9v7y5 drosophila
22	38	65.5	543	10	Q9AWJ8	Q9awj8 oryza sativ
23	38	65.5	832	4	Q96JF7	Q96jf7 homo sapien
24	37	63.8	73	10	Q93VI0	Q93vi0 arabidopsis
25	37	63.8	76	10	Q42337	Q42337 arabidopsis
26	37	63.8	76	10	Q9ZQC9	Q9zqc9 arabidopsis
27	37	63.8	76	10	Q9LFI0	Q9lfi0 arabidopsis
28	37	63.8	175	12	Q9QPN0	Q9qpn0 bovine rota
29	37	63.8	235	10	Q9SRS2	Q9srs2 arabidopsis
30	37	63.8	262	5	Q9VCD6	Q9vcd6 drosophila
31	37	63.8	364	16	Q99V03	Q99v03 staphylococ
32	37	63.8	369	10	O48511	O48511 gracilaria
33	37	63.8	404	4	Q96B18	Q96b18 homo sapien
34	37	63.8	424	4	Q15224	Q15224 homo sapien
35	37	63.8	460	12	Q9WSQ6	Q9wsq6 apple chlor
36	37	63.8	484	16	Q9JXK2	Q9jxk2 neisseria m
37	37	63.8	484	16	Q9JWC7	Q9jwc7 neisseria m
38	37	63.8	767	2	Q9L6V5	Q9l6v5 pseudomonas
39	37	63.8	788	17	O58603	O58603 pyrococcus
40	37	63.8	1311	5	Q9TYW6	Q9tyw6 caenorhabdl
41	37	63.8	1955	4	Q96QU1	Q96qu1 homo sapien
42	37	63.8	1955	4	Q96QT8	Q96qt8 homo sapien
43	36	62.1	150	12	Q99AQ4	Q99aq4 tt virus. o
44	36	62.1	180	4	Q15107	Q15107 homo sapien
45	36	62.1	182	4	Q15108	Q15108 homo sapien

ALIGNMENTS

RESULT 1
Q90Y43 PRELIMINARY; PRT; 115 AA.
AC Q90Y43;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Osteoglossum bicirrhosum (silver arawana).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
OX NCBI_TaxID=109271;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292650; AAK97067.1; -
SQ SEQUENCE 115 AA; 12791 MW; D65FBD7C6F1E4E4D CRC64;

Query Match 79.3%; Score 46; DB 13; Length 115;
Best Local Similarity 80.0%; Pred. No. 0.59;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDNLPPEPRK 11
Db 92 LNNLPPEPRK 101
:::|||||

RESULT 2
Q9V466 PRELIMINARY; PRT; 845 AA.
ID Q9V466
AC Q9V466;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

01-MAR-2001 (TREMELREL. 16, Last annotation update)
BCDNA:LD18761 PROTEIN.
BCDNA:LD18761 OR CG6743.
Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
[1].
SEQUENCE FROM N.A.
STRAIN=BERKELEY;
MEDLINE=20196006; PubMed=10731132;
Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
Abrill J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobaraj C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
The genome sequence of Drosophila melanogaster.;
Science 287:2185-2195(2000).
[2].
SEQUENCE FROM N.A.
Tsang G., Brokstein P., Frise E., Harvey D., Evans-Holm M.,
Lewis S.E., Suh C., Rubin G.M.;
Full Length Drosophila melanogaster cDNA sequence.;
Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
EMBL; AE003629; AAF53002.1; -
EMBL; AF160938; AAD46878.1; -
FlyBase; FBgn0027868; BCDNA:LD18761.
SEQUENCE 845 AA; 97381 MW; 0A75B807A6AFD6B6 CRC64;

Query Match 79.3%; Score 46; DB 5; Length 845;
Best Local Similarity 88.9%; Pred. No. 4.5;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDNLPPE 9
Db 641 SVDNIPPE 649

RESULT 3
013787

ID 013787 PRELIMINARY; PRT; 273 AA.
AC 013787;
DT 01-NOV-1998 (TREMELREL. 08, Created)
DT 01-NOV-1998 (TREMELREL. 08, Last sequence update)
DT 01-NOV-1998 (TREMELREL. 08, Last annotation update)
DE PUTATIVE TRANSLOCATION PROTEIN C17G6.09.
GN SPAC17G6.09.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetaceae;
OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Murphy L., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RL FUNCTION: REQUIRED FOR PREPROTEIN TRANSLOCATION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC
CC RETICULUM (BY SIMILARITY).
CC -!- SIMILARITY: TO YEAST AND Y. LIPOLYTICA SEC62.
DR EMBL; Z99162; CAB16220.1; -
KW Hypothetical protein; Protein transport; Translocation; Transmembrane;
KW Endoplasmic reticulum.
FT DOMAIN 1 132 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 133 153 POTENTIAL.
FT DOMAIN 154 167 LUMENAL (POTENTIAL).
FT TRANSMEM 168 210 POTENTIAL.
FT DOMAIN 211 273 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 273 AA; 31585 MW; EE7A6369530799B3 CRC64;

Query Match 72.4%; Score 42; DB 3; Length 273;
Best Local Similarity 70.0%; Pred. No. 7.7;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VDNLPPE 11
Db 93 VDKLPKQK 102

RESULT 4

Q90Y42 PRELIMINARY; PRT; 114 AA.
ID Q90Y42
AC Q90Y42;
DT 01-DEC-2001 (TREMELREL. 19, Created)
DT 01-DEC-2001 (TREMELREL. 19, Last sequence update)
DT 01-DEC-2001 (TREMELREL. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Pantodon buchholzi (Butterflyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Pantodontidae; Pantodon.
OX NCBI_TaxID=8276;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouti A.A., Irwin D.M., Youson J.H.;
Characterization of variant somatostatin cDNAs from several
osteoglossomorphs: molecular identification and comparative
analysis.;
Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292651; AAK97068.1; -
SQ SEQUENCE 114 AA; 12352 MW; 7E3D44CB6A27B12F CRC64;

Query Match 70.7%; Score 41; DB 13; Length 114;
Best Local Similarity 87.5%; Pred. No. 4.8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 NLPPE 11
Db 93 NVPPPE 100

RESULT 5
 Q9Y3T1 PRELIMINARY; PRT; 242 AA.
 ID Q9Y3T1
 AC Q9Y3T1
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 28.3 KDA PROTEIN (FRAGMENT).
 GN DKFZP564A063.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP TISSUE-BRAIN;
 RC SEQUENCE FROM N.A.
 RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wlennann S.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 JR EMBL; AL050008; CAB43235.1;
 FT NON_TER 1
 SQ SEQUENCE 242 AA; 28250 MW; 84DFC201625C3B1F CRC64;

Query Match 69.0%; Score 40; DB 4; Length 242;
 Best Local Similarity 72.7%; Pred. No. 16;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPrERK 11
 | | | | | | | |
 Db 145 SWDSLPPSERK 155

RESULT 6
 Q98BW0 PRELIMINARY; PRT; 275 AA.
 ID Q98BW0
 AC Q98BW0
 DT 01-OCT-2001 (TReMBLrel. 18, Created)
 DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
 DE MLR5406 PROTEIN.
 GN MLR5406.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AF003006; BAB51862.1;
 DR InterPro; IPR001763; Rhodanese_domain.
 DR SMART; SM00450; RHOD; 1.
 KW Complete proteome.
 SQ SEQUENCE 275 AA; 30613 MW; E12618905C17D6E5 CRC64;

Query Match 69.0%; Score 40; DB 16; Length 275;
 Best Local Similarity 63.6%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SVDNLPPrERK 11
 | | | | | | | |
 Db 110 STDKLPPrDEK 120

RESULT 7
 Q92WH3 PRELIMINARY; PRT; 349 AA.
 ID Q92WH3
 AC Q92WH3
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE PUTATIVE ABC TRANSPORTER ATP-BINDING PROTEIN.
 GN SMB20380.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid pSymb (megaplasmid 2).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396508; PubMed=11481431;
 RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
 RA Vorhoelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
 RA Golding B., Puehler A.;
 RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
 fixing endosymbiont Sinorhizobium meliloti.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL; AL603643; CAC48766.1;
 KW ATP-binding; Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 349 AA; 38734 MW; 17DC3B97F3646393 CRC64;

Query Match 69.0%; Score 40; DB 16; Length 349;
 Best Local Similarity 60.0%; Pred. No. 23;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPrER 10
 : | | | | | : |
 Db 69 AIDNLPPrER 78

RESULT 8
 Q97YY4 PRELIMINARY; PRT; 368 AA.
 ID Q97YY4
 AC Q97YY4
 DT 01-OCT-2001 (TReMBLrel. 18, Created)
 DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
 DE SUGAR ABC TRANSPORTER.
 GN SS01168.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21322296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aways M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moers A., Erauso G., Fletcher C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 DR EMBL; AE006733; AAK41416.1;
 DR InterPro; IPR003593; AAA.
 DR InterPro; IPR003439; ABC_transportr.
 DR InterPro; IPR001687; ATP_GTP_A.
 DR Pfam; PF00005; ABC_tran; 1.
 DR SMART; SM00382; AAA; 1.
 KW Complete proteome.
 SQ SEQUENCE 368 AA; 41884 MW; 6B728206043290F5 CRC64;

Query Match 69.0%; Score 40; DB 17; Length 368;
 Best Local Similarity 60.0%; Pred. No. 24;
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDNLPPEPRK 11
 DB 72 VDDLPPEKDR 81
 ID:|||||

RESULT 9
 Q938D1 PRELIMINARY; PRT; 477 AA.
 AC Q938D1;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 50.9 KDA PROTEIN.
 OS Mycobacterium smegmatis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1772;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-MC2.155;
 RA Derbyshire K.M., Parsons L.M., DeVost J.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY054120; AAL17929.1;
 KW Hypothetical protein.
 SQ SEQUENCE 477 AA; 50879 MW; 6EC7788CA2CDE913 CRC64;

Query Match 69.0%; Score 40; DB 2; Length 477;
 Best Local Similarity 70.0%; Pred. No. 32;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPEPR 10
 DB 384 SIDEPPPTER 393
 ID:|||||

RESULT 10
 Q9N6U5 PRELIMINARY; PRT; 962 AA.
 AC Q9N6U5;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE SU(UR)ES PROTEIN.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA Makunin I.V.;
 RT "Characterization of the Su(UR)ES gene."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ277592; CAB89187.1;
 DR FlyBase; FBgn0025355; Su(UR)ES.
 DR InterPro; IPR000330; SNF2_N.
 DR Pfam; PF00176; SNF2_N; 1.
 SQ SEQUENCE 962 AA; 107624 MW; 5063FB88E875C4F9 CRC64;

Query Match 69.0%; Score 40; DB 5; Length 962;
 Best Local Similarity 60.0%; Pred. No. 65;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPEPR 10

Db 400 ATDNMPPEPR 409

RESULT 11
 Q9VTE2 PRELIMINARY; PRT; 1006 AA.
 AC Q9VTE2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE SUPPRESSOR OF UNDERREPLICATION ES-PROTEIN.
 GN SU(UR)ES OR CG7869.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M., Reese M.G.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003546; AAF50110.2;
 DR FlyBase; FBgn0025355; Su(UR)ES.
 DR InterPro; IPR000330; SNF2_N.
 DR Pfam; PF00176; SNF2_N; 1.
 SQ SEQUENCE 1006 AA; 112638 MW; C281A9520DD24D2A CRC64;

Query Match 69.0%; Score 40; DB 5; Length 1006;
 Best Local Similarity 60.0%; Pred. No. 68;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPEPR 10
 ID:|||||

Db 444 ATDNMPRKR 453

RESULT 12

Q9FQD8 PRELIMINARY; PRT; 219 AA.

ID Q9FQD8

AC Q9FQD8

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE GLUTATHIONE S-TRANSFERASE GST 14 (EC 2.5.1.18) (FRAGMENT)

OS Glycine max (Soybean)

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eudicots I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

OX NCBI_TaxID=3847;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20532786; PubMed=11080288;

RA McGonigle B., Keeler S.J., Lau S.M.C., Koeppe M.K., O'Keefe D.P.;

RT "A Genomics Approach to the Comprehensive Analysis of the Glutathione

RL S-Transferase Gene Family in Soybean and Maize.";

DR EMBL; AF243369; AAG34804.1; -

DR InterPro; IPR004046; GST_C.

DR InterPro; IPR004045; GST_N.

DR Pfam; PF00043; GST_C; 1.

KW Transferase.

FT NON_TER 1

SQ SEQUENCE 219 AA; 25356 MW; 7378601EFB6C7630 CRC64;

Query Match 67.2%; Score 39; DB 10; Length 219;

Best Local Similarity 75.0%; Pred. No. 22;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10

Db 194 DNLPPRDK 201

RESULT 13

Q9FQD8 PRELIMINARY; PRT; 224 AA.

ID Q9FQD8

AC Q9FQD8

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)

DE GLUTATHIONE S-TRANSFERASE GST 20 (EC 2.5.1.18).

OS Glycine max (Soybean)

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eudicots I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

OX NCBI_TaxID=3847;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20532786; PubMed=11080288;

RA McGonigle B., Keeler S.J., Lau S.M.C., Koeppe M.K., O'Keefe D.P.;

RT "A Genomics Approach to the Comprehensive Analysis of the Glutathione

RL S-Transferase Gene Family in Soybean and Maize.";

DR EMBL; AF243375; AAG34810.1; -

DR InterPro; IPR004046; GST_C.

DR InterPro; IPR004045; GST_N.

DR Pfam; PF00043; GST_C; 1.

DR Pfam; PF02798; GST_N; 1.

KW Transferase.

SQ SEQUENCE 224 AA; 26222 MW; 013C0D0C72D630DF CRC64;

Query Match 67.2%; Score 39; DB 10; Length 224;

Best Local Similarity 75.0%; Pred. No. 22;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10

Db 202 DNLPPRDK 209

RESULT 14

Q9RX02 PRELIMINARY; PRT; 444 AA.

ID Q9RX02

AC Q9RX02

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE INTEGRASE, PUTATIVE.

GN DR0513.

OS Deinococcus radiodurans.

OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.

OX NCBI_TaxID=1299;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=RL;

RX MEDLINE=20036896; PubMed=10567266;

RA White O., Eisen J.A., Heideberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,

RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,

RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,

RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,

RA Fraser C.M.;

RT "Genome sequence of the radioresistant bacterium Deinococcus

RT radiodurans R1.";

RL Science 286:1571-1577(1999).

DR EMBL; AE001910; AAF10092.1; -

DR TIGR; DR0513; -

DR InterPro; IPR001994; Cytidylyltransf.

DR InterPro; IPR002104; Phage_integrase.

DR Pfam; PF00589; Phage_integrase; 1.

KW Complete proteome.

SQ SEQUENCE 444 AA; 48293 MW; A885775BD76F0143 CRC64;

Query Match 67.2%; Score 39; DB 16; Length 444;

Best Local Similarity 87.5%; Pred. No. 45;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 DNLPPRR 10

Db 67 DTLPPRR 74

RESULT 15

Q9VAP3 PRELIMINARY; PRT; 796 AA.

ID Q9VAP3

AC Q9VAP3

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE CG11880 PROTEIN.

GN CG11880.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
A Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
A Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003768; AAF56859.1; -;
DR Flybase; FBgn0039637; CG11880.
SQ SEQUENCE 796 AA; 90628 MW; 20961DD889A3EE3B CRC64;

Query Match 67.2%; Score 39; DB 5; Length 796;
Best Local Similarity 66.7%; Pred. No. 82;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 DNLPPRERK 11
Db 786 NNLPPRQR 794

Search completed: June 13, 2002, 12:31:03
Time: 485 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:29:54 ; Search time 18.73 Seconds
(without alignments)
22.740 Million cell updates/sec

Title: US-09-727-739B-18
Perfect score: 58
Sequence: 1 SVDNLPPEPRK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	84.5	125	1 SMS2_LOPAM	P01170 lophius ame
2	45	77.6	73	1 SMS2_PLAFE	P21780 platichthys
3	45	77.6	115	1 SMS2_ONCMY	Q91194 oncorhynchu
4	40	69.0	1061	1 TRC4_ECOLI	P27189 escherichia
5	39	67.2	110	1 YFMT_THETH	P43520 thermus aqu
6	38	65.5	28	1 SMS2_ORENI	P81029 oreochromis
7	37	63.8	76	1 S1FA_ARATH	P42551 arabidopsis
8	37	63.8	78	1 Y32L_HAEIN	Q57534 haemophilus
9	37	63.8	124	1 YIDD_STRCO	O54625 streptomyce
10	37	63.8	250	1 RL7B_SCHPO	P25457 schizosacch
11	37	63.8	251	1 RL7C_SCHPO	O60143 schizosacch
12	37	63.8	339	1 SYFA_DEIRA	Q9rrx8 delnococcus
13	37	63.8	424	1 PSGA_HUMAN	Q15235 homo sapien
14	37	63.8	424	1 PSGC_HUMAN	O75244 homo sapien
15	37	63.8	435	1 PSG6_HUMAN	Q00889 homo sapien
16	37	63.8	641	1 HRPX_PSESY	P41501 pseudomonas
17	37	63.8	1955	1 PC15_HUMAN	Q96qu1 homo sapien
18	36	62.1	246	1 BRM1_HUMAN	Q9hcu9 homo sapien
19	36	62.1	335	1 MAUG_METME	Q50233 methylophil
20	36	62.1	392	1 OR1A_DROME	Q9w5g6 drosophila
21	36	62.1	431	1 VGF_BPAPH	P08767 bacterioph
22	36	62.1	431	1 VGF_BPAPH	Q38041 bacterioph
23	36	62.1	458	1 YNEI_CAEEL	P30640 caenorhabdi
24	36	62.1	1943	1 PC15_MOUSE	Q99pj1 mus musculu
25	35	60.3	325	1 SYFA_THEMEA	Q9wz58 thermotoga
26	35	60.3	345	1 TF2B_YEAST	P29055 saccharomyc
27	35	60.3	359	1 C11D_MOUSE	Q9dd19 mus musculu
28	35	60.3	406	1 CG16_HUMAN	Q9Y305 homo sapien
29	35	60.3	419	1 PSG1_HUMAN	P11464 homo sapien
30	35	60.3	419	1 PSG4_HUMAN	Q00888 homo sapien
31	35	60.3	419	1 PSG7_HUMAN	Q13046 homo sapien
32	35	60.3	426	1 PSGB_HUMAN	Q00887 homo sapien
33	35	60.3	428	1 PSG3_HUMAN	Q16557 homo sapien

34	35	60.3	495	1 OXAL_HUMAN	Q15070 homo sapien
35	35	60.3	591	1 PAK4_HUMAN	O96013 homo sapien
36	35	60.3	714	1 YJL2_YEAST	P47025 saccharomyc
37	35	60.3	825	1 XFP_SCHPO	O74770 schizosacch
38	35	60.3	1221	1 TOP2_TRYBB	P12531 trypanosoma
39	35	60.3	1506	1 PK3G_MOUSE	O70167 mus musculu
40	34.5	59.5	74	1 SMS2_MYOSC	P09876 myoxocephal
41	34	58.6	70	1 S1FA_SPIOL	P42552 spinacia ol
42	34	58.6	155	1 HS21_PHANI	Q01544 pharbitis n
43	34	58.6	222	1 Y066_RICPR	O9ze80 rickettsia
44	34	58.6	249	1 MCRG_METFE	P12973 methanother
45	34	58.6	249	1 RL7A_SCHPO	P17937 schizosacch

ALIGNMENTS

RESULT 1

SMS2_LOPAM

ID SMS2_LOPAM STANDARD; PRT; 125 AA.

AC P01170; Q91066;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr7,Gly10]somatostatin-14].
OS Lophius americanus (American goosefish) (Anglerfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Lophiiformes; Lophiidae; Lophius.
OX NCBI_TaxID=8073;
RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-81052423; PubMed-6107860;

RA Hobart P.M., Crawford R., Shen L., Pictet R., Rutter W.J.;

RT "Cloning and sequence analysis of cDNAs encoding two distinct

somatostatin precursors found in the endocrine pancreas of

RT anglerfish."

RL Nature 288:137-141(1980).

RN [2]

RP PARTIAL SEQUENCE, AND HYDROXYLATION.

RX MEDLINE-87308304; PubMed-2887572;

RA Andrews P.C., Nichols R., Dixon J.E.;

RT "Post-translational processing of preprosomatostatin-II examined

using fast atom bombardment mass spectrometry."

RL J. Biol. Chem. 262:12692-12699(1987).

CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- MISCELLANEOUS: SOMATOSTATIN II MAY HAVE A DIFFERENT DEGREE OF

ACTIVITY OR A DIFFERENT TYPE OF TARGET CELL FROM SOMATOSTATIN I.

CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; V00641; CAA23987.1; -

CC PIR; A01434; RIAFS2.

CC PIR; A27376; A27376.

CC InterPro; IPR004250; Somatostatin.

CC Pfam; PF03002; Somatostatin; 1.

CC Cleavage on pair of basic residues; Hormone; Signal; Hydroxylation;

CC Multigene family.

CC SIGNAL 1 24 POTENTIAL.

CC PROPEP 25 109

CC PEPTIDE 112 125 [TYR7, GLY10]SOMATOSTATIN-14.

CC DISULFID 114 125

CC MOD_RES 120 120 HYDROXYLATION.

CC CONFLICT 77 98 DV -> TG (IN REF. 1).

CC CONFLICT 90 90 G -> E (IN REF. 1).

```
SQ SEQUENCE 125 AA; 14052 MW; 5E14605D7B9A46FE CRC64;

Query Match 84.5%; Score 49; DB 1; Length 125;
Best Local Similarity 81.8%; Pred. No. 0.027;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 11
   I :|||||
Db 101 STNNLPPEPRK 111

RESULT 2
SMS2_PLAPE STANDARD; PRT; 73 AA.
AC P21780;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28;
   [Tyr7,Gly10]somatostatin-14] (Fragments).
E Platicthys flesus (European flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Pleuronectidae; Platicthys.
OX NCBI_TaxID=8260;
RP SEQUENCE.
RC TISSUE=Pancreas; PubMed=2889597;
RX MEDLINE=88029486;
RA Conlon J.M., Davis M.S., Falkner S., Thim L.;
RT "Structural characterization of peptides derived from
RT prosomatostatins I and II isolated from the pancreatic islets of two
RT species of teleostean fish: the daddy sculpin and the flounder.";
RL Eur. J. Biochem. 168:647-652(1987).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
DR PIR; S00169; S00169.
KW Cleavage on pair of basic residues; Hormone; Multigene family.
FT NON_TER 1
FT NON_CONS 10 11
FT NON_CONS 45 46
FT PEPTIDE 46 73 [TYR21, GLY24]SOMATOSTATIN-28.
FT PEPTIDE 60 73 [TYR7, GLY10]SOMATOSTATIN-14.
FT DISULFID 62 73
FT SEQUENCE 73 AA; 7989 MW; CCCBA6B30DCB29BB CRC64;

Query Match 77.6%; Score 45; DB 1; Length 73;
Best Local Similarity 88.9%; Pred. No. 0.079;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPPEPRK 11
   :|||||
Db 51 NNLPPPEPRK 59

RESULT 3
SMS2_ONCMY STANDARD; PRT; 115 AA.
AC Q91194;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28;
   [Tyr7,Gly10]somatostatin-14].
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95354921; PubMed=7628684;
RA Moore C.A., Kittilson J.D., Dahl S.K., Sheridan M.A.;
RT "Isolation and characterization of a cDNA encoding for
RT preprosomatostatin containing [Tyr7, Gly10]-somatostatin-14 from the
RT endocrine pancreas of rainbow trout, Oncorhynchus mykiss.";
RL Gen. Comp. Endocrinol. 98:253-261(1995).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U32471; AAC59695.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
FT SIGNAL 1 18 POTENTIAL.
FT PROPEP 19 87 POTENTIAL.
FT PEPTIDE 88 115 [TYR21, GLY24]SOMATOSTATIN-28 (POTENTIAL).
FT PEPTIDE 102 115 [TYR7, GLY10]SOMATOSTATIN-14.
FT DISULFID 104 115 BY SIMILARITY.
FT SEQUENCE 115 AA; 12963 MW; 520595025FCA6D91 CRC64;

Query Match 77.6%; Score 45; DB 1; Length 115;
Best Local Similarity 88.9%; Pred. No. 0.13;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPPEPRK 11
   :|||||
Db 93 NNLPPPEPRK 101

RESULT 4
TRC4_ECOLI STANDARD; PRT; 1061 AA.
ID TRC4_ECOLI
AC P27189; P27184;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DNA primase trc (EC 2.7.7.-) (Replication primase).
GN TRAC
OS Escherichia coli.
OG Plasmid Incp-beta RP4.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-14 AND 316-333.
RC STRAIN=HB101;
RX MEDLINE=92297959; PubMed=1818755;
RA Miele L., Strack B., Kruff V., Lanka E.;
RT "Gene organization and nucleotide sequence of the primase region of
RT incp plasmids RP4 and R751.";
RL DNA Seq. 2:145-162(1991).
CC -1- FUNCTION: REQUIRED FOR AUTONOMOUS REPLICATION IN ESCHERICHIA COLI.
CC TRANSFERRED INTO THE RECIPIENT CELL DURING BACTERIAL CONJUGATION.
CC CATALYZES THE SYNTHESIS OF SHORT OLIGORIBONUCLEOTIDE PRIMERS WITH
CC CPA... OR PCPA... AT THEIR 5'-TERMINI ON A SINGLE STRANDED
CC TEMPLATE DNA.
CC -1- ALTERNATIVE PRODUCTS: THE TRAC-1 AND TRAC-2 PROTEINS ARE PRODUCED
CC BY THE USE OF ALTERNATIVE INITIATION SITES.
CC -1- SIMILARITY: TO PLASMID R751 TRAC.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X59793; CAA42455.1; -
 DR EMBL; X59793; CAA42456.1; -
 DR PIR; S37667; S37667.
 DR InterPro; IPR002936; Toprim.
 DR Pfam; PF01751; Toprim; 1.
 DR SMART; SM00493; TOPRIM; 1.
 KW DNA replication; transferase; DNA-directed RNA polymerase; Plasmid;
 KW Alternative initiation.
 FT CHAIN 1 1061
 FT CHAIN 316 1061
 FT INIT_MET 316 316
 FT SEQUENCE 1061 AA; 116722 MW; ABC344D2811B9B31 CRC64;

Query Match 69.0%; Score 40; DB 1; Length 1061;
 Best Local Similarity 70.0%; Pred. No. 15;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VDNLPPEPRK 11
 DB 140 IDGLPPEPRK 149

RESULT 5
 YFMT_THETH STANDARD; PRT; 110 AA.
 AC P43520;

DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein in FMT 3'region (Fragment).
 OS Thermus aquaticus (subsp. thermophilus).
 OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
 OX NCBI_TaxID=274;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VK1;
 RX MEDLINE-95050326; PubMed=7961514;

RA Meinel T., Blanquet S.;
 RT "Characterization of the Thermus thermophilus locus encoding peptide
 T deformylase and methionyl-tRNA(fMet) formyltransferase.";
 RL J. Bacteriol. 176:7387-7390(1994).
 CC -1- SIMILARITY: BELONGS TO THE UPF0042 FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X79087; CAA55697.1; -
 DR HSP; P19821; ITAU.
 KW Hypothetical protein; ATP-binding.
 FT NP_BIND 8 15
 FT NON_TER 110 110
 FT SEQUENCE 110 AA; 12316 MW; C4F21341300F9DA6 CRC64;

Query Match 67.2%; Score 39; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDNLPPEPR 8
 DB 140 IDGLPPEPR 149

DB 29 VDNLPPEPR 35

RESULT 6

SMS2_ORENI STANDARD; PRT; 28 AA.
 AC P81029;

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28;
 DE [Tyr7,Gly10]somatostatin-14] (Fragment).
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
 OC Cichlidae; Oreochromis.
 OX NCBI_TaxID=8128;
 RN [1]
 RP SEQUENCE.

RX MEDLINE-95384941; PubMed=7656183;
 RA Nguyen T.M., Wright J.R. Jr., Nielsen P.F., Conlon J.M.;
 RT "Characterization of the pancreatic hormones from the Brockmann body
 of the tilapia: implications for islet xenograft studies.";
 RL Comp. Biochem. Physiol. 111C:33-44(1995).

CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
 DR InterPro; IPR004250; Somatostatin.
 DR Pfam; PF03002; Somatostatin; 1.
 KW Cleavage on pair of basic residues; Hormone; Multigene family.

FT NON_TER 1 28
 FT PEPTIDE 1 28 [TYR21, GLY24]SOMATOSTATIN-28.
 FT PEPTIDE 15 28 [TYR7, GLY10]SOMATOSTATIN-14.
 FT DISULFID 17 28
 FT SEQUENCE 28 AA; 3155 MW; 47C049F4866EF4AC CRC64;

Query Match 65.5%; Score 38; DB 1; Length 28;
 Best Local Similarity 66.7%; Pred. No. 0.51;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPPEPRK 11
 DB 6 NSIPPEPRK 14

RESULT 7

SIFA_ARATH STANDARD; PRT; 76 AA.
 AC P42551;

DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA binding protein SIFA.
 GN SIFA.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-CV. COLUMBIA;
 RX MEDLINE-95148729; PubMed=7846151;
 RA Newman T., de Bruijn F.J., Green P., Keegstra K., Kende H.,
 RA McIntosh L., Ohlrogge J.B., Raikhel N., Somerville S., Thomashow M.F.,
 RA Retzel E., Somerville C.R.;

RT "Genes galore: a summary of methods for accessing results from large-
 RT scale partial sequencing of anonymous Arabidopsis cDNA clones.";
 RL Plant Physiol. 106:1241-1255(1994).
 CC -1- FUNCTION: DNA-BINDING PROTEIN THAT SPECIFICALLY RECOGNIZES A
 CC NEGATIVE ELEMENT (SIF) WITHIN THE RPS1 PROMOTER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: STRONG, TO SIFA IN OTHER PLANTS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; T20524; -; NOT_ANNOTATED_CDS.
 CC Transcription regulation; Repressor; DNA-binding; Nuclear protein.
 CC SEQUENCE 76 AA; 8252 MW; F05E3D5950F8A8AF CRC64;
 CC -----
 CC Query Match 63.8%; Score 37; DB 1; Length 76;
 CC Best Local Similarity 75.0%; Pred. No. 2.5;
 CC Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC Y 4 NLPPIRERK 11
 CC |||||:
 CC 48 NLPPIRKKK 55
 CC -----
 CC RESULT 8
 CC Y321_HAEIN STANDARD; PRT; 78 AA.
 CC ID Y321_HAEIN STANDARD; PRT; 78 AA.
 CC AC Q57534;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Protein HI0321.
 CC GN HI0321.
 CC OS Haemophilus influenzae.
 CC OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC OC Haemophilus.
 CC OX NCBI_TaxID=727;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN-RD / KW20 / ATCC 51907;
 CC MEDLINE-95350630; PubMed-7542800;
 CC RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 CC Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 CC McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 CC Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 CC Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 CC Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 CC Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 CC Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 CC Venter J.C.;
 CC "Whole-genome random sequencing and assembly of Haemophilus influenzae
 CC Rd.";
 CC Science 269:496-512(1995).
 CC [2]
 CC IDENTIFICATION BY MASS SPECTROMETRY.
 CC MEDLINE-20137488; PubMed-10675023;
 CC RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
 CC Gray C., Fountoulakis M.;
 CC "Two-dimensional map of the proteome of Haemophilus influenzae.";
 CC Electrophoresis 21:411-429(2000).
 CC -1- SIMILARITY: SOME, TO B.NODOSUS VIRULENCE-ASSOCIATED PROTEIN B
 CC (VAPB).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U32717; AAC21984.1; -

DR TIGR; HI0321; -
 KW Complete proteome.
 SQ SEQUENCE 78 AA; 9037 MW; 468D048ED7C5C720 CRC64;

 Query Match 63.8%; Score 37; DB 1; Length 78;
 Best Local Similarity 66.7%; Pred. No. 2.6;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

 QY 3 DNLPPRERK 11
 |:||||:
 Db 68 DDLPPQERE 76

 RESULT 9
 YIDD_STRCO STANDARD; PRT; 124 AA.
 ID YIDD_STRCO STANDARD; PRT; 124 AA.
 AC Q54625;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 13.6 kDa protein STH24.04.
 GN STH24.04.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinomycetaceae; Streptomyces.
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2) / M145;
 RX MEDLINE-99014240; PubMed-9795152;
 RA Gal-Mor O., Borovok I., Av-Gay Y., Cohen G., Aharonowitz Y.;
 RT "Gene organization in the trxA/B-oric region of the Streptomyces
 RT coelicolor chromosome and comparison with other eubacteria.";
 RL Gene 217:83-90(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Kim H.-D., Calcutt M.J., Schmidt F.J., Chater K.F.;
 RT "Partitioning of the linear chromosome during sporulation of
 RT Streptomyces coelicolor A3(2) involves an Oric-linked parAB locus.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Oliver K., Harris D., James K.D., Parkhill J., Barrell B.G.,
 RA Rajandream M.A.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE UPF0161 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Y16311; CAA76173.1; -
 DR EMBL; AL049826; CAB42698.1; -
 DR EMBL; AF187159; AAF16009.1; -
 DR EMBL; AF031590; AAC03488.1; -
 DR InterPro; IPR002696; DUF37;
 DR Pfam; PF01809; DUF37; 1;
 DR ProDom; PD004225; DUF37; 1;
 KW Hypothetical protein.
 SQ SEQUENCE 124 AA; 13599 MW; 8BB3AD1786D8075E CRC64;

 Query Match 63.8%; Score 37; DB 1; Length 124;
 Best Local Similarity 66.7%; Pred. No. 4.4;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AE002066; AAF11900.1; -
DR HSSP; P27001; 1PYS.
DR TIGR; DR2354; -
DR InterPro; IPR002106; AA_trna_ligase_II.
DR InterPro; IPR004188; Phe_trna_synt_N.
DR InterPro; IPR002319; trna_synt_2d.
DR Pfam; PF02912; Phe_trna_synt_N; 1.
DR Pfam; PF01409; trna_synt_2d; 1.
DR PROSITE; PS00179; AA_trna_ligase_II_1; 1.
DR PROSITE; PS00339; AA_trna_ligase_II_2; 1.
KW Aminocyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 339 AA; 37549 MW; FA02BCD1C83A5BB5 CRC64;

Query Match 63.8%; Score 37; DB 1; Length 339;
Best Local Similarity 63.6%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SVDNLPPEERK 11
DB 38 SLGKLPPEERK 48

RESULT 13
PSGA_HUMAN STANDARD; PRT; 424 AA.
ID Q15235;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pregnancy-specific beta-1-glycoprotein 10 precursor (PSBG-10).
GN PSG10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=91104939; PubMed=2271648;
RA Barnett T.R., Pickle W. II, Elting J.J.;
RT "Characterization of two new members of the pregnancy-specific beta
1-glycoprotein family from the myeloid cell line KG-1 and suggestion
of two distinct classes of transcription unit.";
RL Biochemistry 29:10213-10218(1990).
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
CC -1- DEVELOPMENTAL STAGE: PSBG ARE PRODUCED IN HIGH QUANTITY DURING
PREGNANCY.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CEA
SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17098; CAA34957.1; -
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003600; Ig_Llike.
DR Pfam; PF00047; Ig; 4.

DR SMART; SM00410; IG_Like; 3.
DR SMART; SM00408; IGC2; 1.
KW Immunoglobulin domain; Glycoprotein; Signal; Repeat; Multigene family.
FT SIGNAL 1 34 BY SIMILARITY.
FT CHAIN 35 424 PREGNANCY-SPECIFIC BETA-1-GLYCOPROTEIN
FT 10.
FT DOMAIN 35 143 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 161 223 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 254 316 IG-LIKE C2-TYPE DOMAIN 2.
FT DOMAIN 346 400 IG-LIKE C2-TYPE DOMAIN 3.
FT SITE 126 128 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 168 216 PROBABLE.
FT DISULFID 261 309 PROBABLE.
FT DISULFID 353 393 PROBABLE.
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 424 AA; 47676 MW; 0B7726090B31B92E CRC64;

Query Match 63.8%; Score 37; DB 1; Length 424;
Best Local Similarity 54.5%; Pred. No. 18;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPEERK 11
DB 244 TINNLNPREKK 254

RESULT 14
PSGC_HUMAN STANDARD; PRT; 424 AA.
ID O75244;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pregnancy-specific beta-1-glycoprotein 12 precursor (PSBG-12).
GN PSG12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K.J., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Barnes J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Trankhelm M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
CC -1- DEVELOPMENTAL STAGE: PSBG ARE PRODUCED IN HIGH QUANTITY DURING
PREGNANCY.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CEA
SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

FT	CARBOHYD	386	386	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	155	155	L -> F (IN DBSNP:1058674).
FT				/FTid=VAR_011723.
FT	VARIANT	165	165	R -> L (IN DBSNP:1058680).
FT				/FTid=VAR_011724.
FT	VARIANT	167	167	I -> T (IN DBSNP:1065505).
FT				/FTid=VAR_011725.
FT	VARIANT	180	180	L -> W (IN DBSNP:1065507).
FT				/FTid=VAR_011726.
FT	VARIANT	181	181	L -> M (IN DBSNP:1065508).
FT				/FTid=VAR_011727.
FT	VARIANT	185	185	N -> S (IN DBSNP:1065509).
FT				/FTid=VAR_011728.
FT	VARIANT	191	191	R -> S (IN DBSNP:1058688).
FT				/FTid=VAR_011729.
FT	VARIANT	196	196	K -> E (IN DBSNP:1065511).
FT				/FTid=VAR_011730.
FT	VARIANT	253	253	K -> N (IN DBSNP:1065513).
FT				/FTid=VAR_011731.
FT	VARIANT	257	257	L -> S (IN DBSNP:1058710).
FT				/FTid=VAR_011732.
FT	VARIANT	258	258	A -> T (IN DBSNP:1065515).
FT				/FTid=VAR_011733.
FT	VARIANT	344	344	R -> H (IN DBSNP:1065519).
FT				/FTid=VAR_011734.
FT	VARIANT	404	404	I -> S (IN DBSNP:1065525).
FT				/FTid=VAR_011735.
SQ	SEQUENCE	435 AA;	48813 MW;	D94D4B0884327C60 CRC64;

Query Match 63.8%; Score 37; DB 1; Length 435;
Best Local Similarity 54.5%; Pred. NO. 19;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SVDNLPPEK 11
Db 244 TNNLPPEK 254

Search completed: June 13, 2002, 12:29:54
Job time: 446 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:22:56 ; Search time 28.68 Seconds
(without alignments)
9.368 Million cell updates/sec

Title: US-09-727-739B-18
Perfect score: 58
Sequence: 1 SVDNLPPTPRK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000.

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	67.2	219	4 US-09-247-373B-52	Sequence 52, Appl
2	39	67.2	224	4 US-09-247-373B-34	Sequence 34, Appl
3	37	63.8	424	6 5169835-6	Patent No. 5169835
4	36	62.1	2496	4 US-09-125-028-2	Sequence 2, Appl
5	36	62.1	2958	4 US-08-894-344C-2	Sequence 2, Appl
6	35	60.3	230	6 5169835-13	Patent No. 5169835
7	35	60.3	419	6 5169835-2	Patent No. 5169835
8	35	60.3	591	3 US-09-082-737-2	Sequence 2, Appl
9	34	58.6	223	3 US-08-857-534-12	Sequence 12, Appl
10	34	58.6	223	5 PCT-US95-04971-12	Sequence 12, Appl
11	34	58.6	229	4 US-09-247-373B-48	Sequence 48, Appl
12	34	58.6	361	1 US-08-258-261B-3	Sequence 3, Appl
13	34	58.6	361	1 US-08-456-837-3	Sequence 3, Appl
14	34	58.6	361	1 US-08-457-342-3	Sequence 3, Appl
15	34	58.6	361	1 US-08-457-646A-3	Sequence 3, Appl
16	34	58.6	361	1 US-08-458-076A-3	Sequence 3, Appl
17	34	58.6	361	1 US-08-457-335A-3	Sequence 3, Appl
18	34	58.6	361	2 US-08-723-214-3	Sequence 3, Appl
19	34	58.6	361	3 US-09-028-934-3	Sequence 3, Appl
20	33	56.9	265	2 US-08-970-133-1	Sequence 1, Appl
21	33	56.9	265	4 US-09-294-545-1	Sequence 1, Appl
22	33	56.9	345	1 US-08-031-148-4	Sequence 4, Appl
23	33	56.9	345	3 US-08-415-838-4	Sequence 4, Appl
24	33	56.9	415	2 US-08-576-626A-52	Sequence 52, Appl
25	33	56.9	683	1 US-07-878-960-2	Sequence 2, Appl
26	33	56.9	855	1 US-08-344-536-2	Sequence 2, Appl
27	33	56.9	855	1 US-08-344-536-5	Sequence 5, Appl

28	33	56.9	855	3 US-08-920-562-2	Sequence 2, Appl
29	33	56.9	855	3 US-08-920-562-5	Sequence 5, Appl
30	33	56.9	2544	2 US-08-576-626A-32	Sequence 32, Appl
31	32	55.2	275	4 US-09-147-915-4	Sequence 4, Appl
32	32	55.2	296	3 US-08-986-769-2	Sequence 2, Appl
33	32	55.2	361	3 US-09-028-934-30	Sequence 30, Appl
34	32	55.2	449	1 US-08-624-663A-2	Sequence 2, Appl
35	32	55.2	449	2 US-08-974-565C-1	Sequence 1, Appl
36	32	55.2	449	3 US-09-255-748-1	Sequence 1, Appl
37	32	55.2	470	4 US-09-004-838-44	Sequence 44, Appl
38	32	55.2	472	4 US-09-004-838-43	Sequence 43, Appl
39	32	55.2	560	2 US-08-643-034A-2	Sequence 2, Appl
40	32	55.2	560	3 US-08-648-650A-2	Sequence 2, Appl
41	32	55.2	600	3 US-08-904-871-3	Sequence 3, Appl
42	32	55.2	713	2 US-08-987-466-3	Sequence 3, Appl
43	32	55.2	713	2 US-08-974-565C-5	Sequence 5, Appl
44	32	55.2	713	3 US-09-255-748-5	Sequence 5, Appl
45	32	55.2	713	4 US-09-240-359-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-247-373B-52
; Sequence 52, Application US/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1108-A
; CURRENT APPLICATION NUMBER: US/09/247,373B
; CURRENT FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 08/924,747
; PRIOR FILING DATE: 1997-09-05
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 52
; LENGTH: 219
; TYPE: PRT
; ORGANISM: SOYBEAN
US-09-247-373B-52

Query Match 67.2% ; Score 39; DB 4; Length 219;
Best Local Similarity 75.0% ; Pred. No. 8.5;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 DNLPPRR 10
Db 194 DNLPPRDK 201

RESULT 2
US-09-247-373B-34
; Sequence 34, Application US/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1108-A
; CURRENT APPLICATION NUMBER: US/09/247,373B
; CURRENT FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 08/924,747
; PRIOR FILING DATE: 1997-09-05
; NUMBER OF SEQ ID NOS: 56
; SEQ ID NO 34
; LENGTH: 224
; TYPE: PRT
; ORGANISM: SOYBEAN

US-09-247-373B-34

Query Match 67.2%; Score 39; DB 4; Length 224;
Best Local Similarity 75.0%; Pred. No. 8.7;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10
| | | | | | | | | |
Db 202 DNLPPRDK 209

RESULT 3
5169835-6
; Patent No. 5169835
; APPLICANT: WAI-YEE, CHAN
; TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/390,409
; FILING DATE: 07-AUG-1989
; SEQ ID NO: 6
; LENGTH: 424
5169835-6

Query Match 63.8%; Score 37; DB 6; Length 424;
Best Local Similarity 54.5%; Pred. No. 39;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPERRK 11
: : : | | | | | : |
Db 244 TINNLNPREKK 254

RESULT 4
US-09-125-028-2
; Sequence 2, Application US/09125028A
; Patent No. 6190707
; GENERAL INFORMATION:
; APPLICANT: WADOUX, Isabelle
; APPLICANT: COLAVIZZA, Didier
; APPLICANT: LOIEZ, Annie
; TITLE OF INVENTION: NOVEL COLD-SENSITIVE BREAD-MAKING YEASTS.
; FILE REFERENCE: levure sensible froid
; CURRENT APPLICATION NUMBER: US/09/125,028A
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: PCT/FR97/00254
; EARLIER FILING DATE: 1997-02-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2496
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-125-028-2

Query Match 62.1%; Score 36; DB 4; Length 2496;
Best Local Similarity 66.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VDNLPERR 10
: | | | | | | | | | |
Db 257 IDVLPPKER 265

RESULT 5
US-08-894-344C-2
; Sequence 2, Application US/08894344C
; Patent No. 6172196
; GENERAL INFORMATION:
; APPLICANT: KAWASAKI, Hideki

APPLICANT: TOKAI, Masaya
APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: OUCHI, KOZO
TITLE OF INVENTION: DNA ENCODING PROTEIN COMPLEMENTING
TITLE OF INVENTION: YEAST
TITLE OF INVENTION: LOW TEMPERATURE-SENSITIVE FERMENTABILITY
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 1440 Kb storage.
COMPUTER: IBM PS/V
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,344C
FILING DATE: 15-AUGUST-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP343700/95
FILING DATE: 28-DECEMBER-1995
APPLICATION NUMBER: PCT/JP96/03862
FILING DATE: 27-DECEMBER-1996
ATTORNEY/AGENT INFORMATION:
NAME: Perry, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2958 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Saccharomyces cerevisiae
STRAIN: X2180-1B
US-08-894-344C-2

Query Match 62.1%; Score 36; DB 4; Length 2958;
Best Local Similarity 66.7%; Pred. No. 4.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VDNLPERR 10
: | | | | | | | | | |
Db 257 IDVLPPKER 265

RESULT 6
5169835-13
; Patent No. 5169835
; APPLICANT: WAI-YEE, CHAN
; TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/390,409
; FILING DATE: 07-AUG-1989
; SEQ ID NO: 13
; LENGTH: 230
5169835-13

Query Match 60.3%; Score 35; DB 6; Length 230;
Best Local Similarity 54.5%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPERRK 11

Db 101 TNNLPPRENK 111
:::11 1111 1

RESULT 7
5169835-2
; Patent No. 5169835
; APPLICANT: WAI-YEE, CHAN
; TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/390,409
; FILING DATE: 07-AUG-1989
; SEQ ID NO: 2:
; LENGTH: 419
5169835-2

Query Match 60.3%; Score 35; DB 6; Length 419;
Best Local Similarity 54.5%; Pred. No. 88;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 SVDNLPPEPRK 11
:::11 1111 1

Db 245 TNNLPPRENK 255

RESULT 8
US-09-082-737-2
; Sequence 2, Application US/09082737
; Patent No. 6013500
; GENERAL INFORMATION:
; APPLICANT: Minden, Audrey
; TITLE OF INVENTION: PAK4; A No. 6013500el Gene Encoding A Serine/
; TITLE OF INVENTION: Threonine Kinase
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11230
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/082,737
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/55311
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-082-737-2

Query Match 60.3%; Score 35; DB 3; Length 591;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 DNLPPR 8

Db 529 DNLPPR 534
111111

RESULT 9
US-08-857-534-12
; Sequence 12, Application US/08857534
; Patent No. 6087170
; GENERAL INFORMATION:
; APPLICANT: George W. Kemble
; TITLE OF INVENTION: A No. 6087170el VZV Gene, Mutant VZV and Immunogenic
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/857,534
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/235,406
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Luann Cseri
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: AVIR-004/000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-843-5165
; TELEFAX: 415-857-0663
; TELEX: 380816 CooleyPA
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: Protein
US-08-857-534-12

Query Match 58.6%; Score 34; DB 3; Length 223;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PPRERK 11
111111

Db 44 PPRERK 49

RESULT 10
PCT-US95-04971-12
; Sequence 12, Application PC/TUS9504971
; GENERAL INFORMATION:
; APPLICANT: George W. Kemble
; TITLE OF INVENTION: A Novel VZV Gene, Mutant VZV and Immunogenic
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04971
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,406
; FILING DATE: APRIL 28, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Luann Cseri
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: AVIR-004/00WO
; TELEPHONE: 415-843-5165
; TELEFAX: 415-857-0663
; TELEX: 380816 CooleyPA
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: Protein
; PCT-US95-04971-12

```

```

Query Match 58.6%; Score 34; DB 5; Length 223;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 6 PPRERK 11
Db 44 PPRERK 49

```

```

RESULT 11
US-09-247-373B-48
; Sequence 48, Application US/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1108-A
; CURRENT APPLICATION NUMBER: US/09/247,373B
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 08/924,747
; PRIOR FILING DATE: 1997-09-05
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 48
; LENGTH: 229
; TYPE: PRT
; ORGANISM: SOYBEAN
; US-09-247-373B-48

```

```

Query Match 58.6%; Score 34; DB 4; Length 229;
Best Local Similarity 62.5%; Pred. No. 70;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 DNLPPRR 10
Db 200 DSLPPRDK 207

```

```

RESULT 12
US-08-258-261B-3
; Sequence 3, Application US/08258261B
; Patent No. 5639949

```

```

; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/258,261B
; FILING DATE: 08-JUN-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-258-261B-3

```

```

Query Match 58.6%; Score 34; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 NLPPE 9
Db 107 NLPPE 112

```

```

RESULT 13
US-08-456-837-3
; Sequence 3, Application US/08456837
; Patent No. 5643774
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances

```

NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,837
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 361 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-456-837-3

Query Match 58.6%; Score 34; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 NLPPRE 9
Db 107 NLPPRE 112

RESULT 14
US-08-457-342-3
Sequence 3, Application US/08457342
Patent No. 5662898
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,342
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 361 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-457-342-3

Query Match 58.6%; Score 34; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 NLPPRE 9
Db 107 NLPPRE 112

RESULT 15
US-08-457-646A-3
Sequence 3, Application US/08457646A
Patent No. 5679560
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,646A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995

```

; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-457-646A-3

```

```

Query Match      58.6%; Score 34; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 NLPPE 9
Db 107 NLPPE 112

```

Search completed: June 13, 2002, 12:22:56
Job time: 194 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:22:21 ; Search time 72.45 seconds
(without alignments)
16.864 Million cell updates/sec

Title: US-09-727-739B-18
Perfect score: 58
Sequence: 1 SVDNLPPEPRK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	111	AAU07668	Rainbow trout prep
2	49	84.5	28	AAP61714	Somatostatin-28 an
3	49	84.5	28	AAP90989	Analogue of angler
4	49	84.5	125	AAP20029	Sequence of prepro
5	46	79.3	845	ABG62651	Drosophila melanog
6	45	77.6	115	AAU07667	Rainbow trout prep
7	40	69.0	290	AAM38940	Human polypeptide
8	40	69.0	1006	ABG63556	Drosophila melanog
9	39	67.2	219	AAU07849	Amino acid sequenc
10	39	67.2	224	AAU07840	Amino acid sequenc
11	39	67.2	244	AAU07843	S. epidermidis ope

12	39	67.2	370	22	ABB52820	Escherichia coli p
13	39	67.2	796	22	ABB69479	Drosophila melanog
14	38	65.5	33	19	AAW79680	Synthetic CS198 de
15	38	65.5	33	22	AAE07323	Human CS 198 pepti
16	38	65.5	175	22	AAB93866	Human protein sequ
17	38	65.5	197	22	AAU17725	Novel human respir
18	38	65.5	281	21	AAU42000	Human ORFX ORF1764
19	38	65.5	486	22	ABB59435	Drosophila melanog
20	38	65.5	490	22	ABB70830	Drosophila melanog
21	37	63.8	56	22	AAM89715	Human immune/haema
22	37	63.8	73	21	AAG28322	Arabidopsis thalia
23	37	63.8	76	21	AAG04246	Arabidopsis thalia
24	37	63.8	76	21	AAG10791	Arabidopsis thalia
25	37	63.8	76	21	AAG42736	Arabidopsis thalia
26	37	63.8	76	21	AAG54001	Arabidopsis thalia
27	37	63.8	76	22	AAE02520	Arabidopsis thalia
28	37	63.8	110	21	AAG04245	Arabidopsis thalia
29	37	63.8	110	21	AAG54000	Arabidopsis thalia
30	37	63.8	262	22	ABB65208	Drosophila melanog
31	37	63.8	392	22	ABG18522	Novel human diagno
32	37	63.8	424	10	AAP93998	Amino acid sequenc
33	37	63.8	424	11	AAR06430	SPI-like protein e
34	37	63.8	497	22	ABG21065	Novel human diagno
35	36	62.1	76	21	AAG59881	Arabidopsis thalia
36	36	62.1	228	22	AAM39970	Human polypeptide
37	36	62.1	245	22	AAM41756	Human polypeptide
38	36	62.1	268	22	AAG82368	S. epidermidis ope
39	36	62.1	287	22	ABB58433	Drosophila melanog
40	36	62.1	290	21	AAB42044	Human ORFX ORF1808
41	36	62.1	304	22	AAM39860	Human polypeptide
42	36	62.1	315	22	AAM40726	Human polypeptide
43	36	62.1	381	21	AAU26450	Drosophila melanog
44	36	62.1	381	22	ABB66532	Drosophila melanog
45	36	62.1	392	21	AAB20945	Drosophila odorant

ALIGNMENTS

RESULT 1
AAU07668
ID AAU07668 standard; Protein; 111 AA.
XX
AC AAU07668;
XX
DT 04-DEC-2001 (first entry)
XX
DE Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.
XX
KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
KW PPSS-II'; PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;
KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;
KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;
KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.
XX
OS Oncorhynchus mykiss.

Key	Location/Qualifiers
Peptide	1..25
Protein	/note= "Signal peptide"
Protein	/note= 1..86
Protein	/note= "PPSS-II'' pre-sequence"
Protein	26..111
Protein	/note= "Mature PPSS-II''"
Protein	/note= 87..97
Protein	/note= "PPSS-II'' pro-sequence"
Protein	87..111
Protein	/note= "Prosomatostatin II''"
Cleavage-site	96..97
	/note= "Dibasic cleavage site"

FT Peptide 98..111 /note="SS-14 variant peptide"
FT CA2325169-A1.
PN 03-JUN-2001.
XX 01-DEC-2000; 2000CA-2325169.
XX 03-DEC-1999; 99US-0168934.
XX (NDSU-) NDSU RES FOUND.
PA Sheridan MA, Moore CA, Kittelson JD;
PI WPI; 2001-425997/46.
XX N-PSDB; AAS12935.

New somatostatin polypeptides derived from Oncorhynchus mykiss, useful for treating diabetes mellitus, acromegaly, gastrinoma, acquired immunodeficiency syndrome and neurological disorders -

Claim 1; Fig 3; 52pp; English.

The invention relates to an Oncorhynchus mykiss somatostatin polypeptide containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of preprosomatostatin II (PPSS-II). The protein sequences and their associated polynucleotides are useful for identifying modified somatostatin polypeptide which functions as a somatostatin agonist useful for research, therapeutics or diagnostics, including medical and veterinary applications. The wild-type somatostatin and its modified version are useful for treating hypersecretion from endocrine tumours in the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g. gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage through their effects on cell proliferation and apoptosis and as adjuncts in the treatment of diabetes mellitus via inhibition of growth hormone and glucagon. In addition, dysfunctional somatostatin secretion is associated with acquired immunodeficiency syndrome (AIDS) and various neurological disorders (e.g. epilepsy, Alzheimer's disease and Huntington's disease) and somatostatin antagonists are effective in the treatment of such conditions. Nucleic acids encoding the polypeptides are useful in gene therapy and fusion peptides can be targeted to neoplasms and their metastases, inhibiting the release of their secretory products. This sequence represents O. Mykiss PPSS-II', protein.
Note: The features for this sequence are specifically claimed in the specification.

Sequence 111 AA;

Query Match 100.0%; Score 58; DB 22; Length 111;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDNLPPrERK 11
Db 87 svdnlpprerk 97
|||||

RESULT 2
AAP61714
ID AAP61714 standard; Protein; 28 AA.
XX
AC AAP61714;
XX

DT 28-JUL-1991 (first entry)
XX
DE Somatostatin-28 analogue.
XX
KW Somatostatin-28; insulin-selective; insulinoma.
XX
OS Synthetic.
XX

FH Key Location/Qualifiers
FT Misc-difference 23 /label= Hyl, Lys
XX EP173527-A.
XX 05-MAR-1986.
XX 16-AUG-1985; 85EP-0305867.
XX 31-AUG-1984; 84US-0646610.
XX (SALK) SALK INST FOR BIOL STUD.
XX Spiess J, Noe BD;
XX WPI; 1986-063363/10.
XX Angler fish somatostatin-28 and analogue and fragment - useful in inhibiting insulin secretion in insulinoma.
XX Claim 4; Page 18; 19pp; English.
XX The protein sequence is an insulin-selective analogue of anglerfish somatostatin-28, which is more potent than somatostatin-14 or somatostatin-28 in inhibiting insulin secretion for treatment of insulinoma.
XX Sequence 28 AA;

Query Match 84.5%; Score 49; DB 7; Length 28;
Best Local Similarity 81.8%; Pred. No. 0.048;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPrERK 11
Db 4 stnlprrerk 14
|:|||||

RESULT 3
AAP90989
ID AAP90989 standard; peptide; 28 AA.
XX
AC AAP90989;
XX

DT 08-JUN-1990 (first entry)
XX
DE Analogue of anglerfish somatostatin 28.
XX

KW Somatostatin 28; SS-28; analogue; insulin secretion inhibitor; insulinoma; gastric acid secretion; thermoregulation.
XX Anglerfish.

FH Key Location/Qualifiers
FT Disulfide-bond 17 /note="Bonded to Cys-28"
FT Disulfide-bond 28 /note="Bonded to Cys-17"
FT Misc-difference 23 /label=Lys, Hyl
FT Region 15..28 /note="Also claimed"

XX US4816438-A.
PN 28-MAR-1989.
XX
XX 01-APR-1987; 87US-0033295.
XX 01-APR-1987; 87US-0033295.
XX 31-AUG-1984; 84US-0646610.
PR

XX PA (SALK) SALK INST FOR BIOL STUD.
XX PI Spiess J, Noe BD;
XX DR WPI; 1989-113910/15.
XX XX Angler fish somatostatin-28 and fragments -
PT useful in inhibiting insulin secretion and insulinoma
PT XX
PS Claim 1; page 65; 8pp; English.
XX It is called ass-28 because it is an analogue of anglerfish somatostatin
CC (SS-28). It is more potent than either somatostatin 14 (SS-14) or SS-28
CC at inhibiting insulin secretion for the treatment of insulinoma. The
CC 14-residue C-terminal peptide is also claimed (ass-14). ass-14 is useful
CC for inhibiting insulin secretion by the pancreas. ass-28 and ass-28 may
CC be useful for decreasing gastric acid secretion and influencing
CC thermoregulation. Their reduced linear forms, wherein the disulphide
CC bridge is not present and is replaced by H, is also claimed.
XX
XX Sequence 28 AA;

Query Match 84.5%; Score 49; DB 10; Length 28;
Best Local Similarity 81.8%; Pred. No. 0.048;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 11
Db 4 stnnlpprerk 14

RESULT 4
AAP20029
ID AAP20029 standard; Protein; 125 AA.
AC AAP20029;
XX 14-AUG-1992 (first entry)
XX Sequence of preprosomatostatin-2 encoded on pLas2.
DE Somatostatin; growth hormone; peptide hormone; secretion.
XX Somatostatin; growth hormone; peptide hormone; secretion.
XX Key Location/Qualifiers
FH Protein 112..125
TT /label= Somatostatin II
XX
PN EP46669-A.
XX
PD 03-MAR-1982.
XX
PF 21-AUG-1981; 81EP-0303825.
XX
PR 25-AUG-1980; 80US-0181046.
XX
PA (REGC) UNIV OF CALIFORNIA;
XX Hobart P, Crawford R, Pictet RL, Rutter WJ;
PI WPI; 1982-18113E/10.
XX N-PSDB; AAN20034.
DR New somatostatin and precursors - produced by transformed
XX microorganisms
PT Example; Fig 3; 50pp; English.
XX The inventors claim preprosomatostatin-1, prosomatostatin-1,
CC preprosomatostatin-2, prosomatostatin-2 and somatostatin-2; and DNA
CC encoding them. The translation of somatostatin mRNA yields a
CC precursor (prepro S1) containing a signal peptide which may be

CC released during the transit into the endoplasmic reticulum, and the
CC resultant precursor (pro S1) is subsequently cleaved to yield S1
CC itself. The prepeptide portion of prepro S1 is probably about 20-25
CC bases long. Translation of pLas2 predicts the sequence of a 125 AA
CC peptide which surprisingly contains a 14 AA sequence at its carboxy
CC terminus which differs from S1 by only 2 AAs, and is termed
CC Somatostatin 2 (S2).
XX
XX Sequence 125 AA;

Query Match 84.5%; Score 49; DB 3; Length 125;
Best Local Similarity 81.8%; Pred. No. 0.23;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 11
Db 101 stnnlpprerk 111

RESULT 5
ABB62651
ID ABB62651 standard; Protein; 845 AA.

XX AC ABB62651;
XX 26-MAR-2002 (first entry)
DT Drosophila melanogaster polypeptide SEQ ID NO 14745.
DE Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX WO200171042-A2.
PN 27-SEP-2001.
PD 23-MAR-2001; 2001WO-US09231.
PF 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
PR (PEKE) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
DR N-PSDB; ABL06754.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Disclosure; SEQ ID NO 14745; 21pp + Sequence Listing; English.
PS The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 845 AA;

Query Match 79.3%; Score 46; DB 22; Length 845;

Best Local Similarity 88.9%; Pred. No. 5.9; Mismatches 1; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVDNLPPE 9
Db 641 svdnppre 649
RESULT 6
ID AAU07667 standard; Protein; 115 AA;
XX
AC AAU07667;
DT 04-DEC-2001 (first entry)
XX
DE Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.
XX
KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
KW PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
W carcinoid syndrome; cell proliferation; apoptosis; growth hormone;
KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;
KW neoplasia; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;
KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.
XX
OS Oncorhynchus mykiss.
XX
FH Key Location/Qualifiers
FT Peptide 1..25 /note= "Signal peptide"
FT Protein 1..87 /note= "PPSS-II' pre-sequence"
FT Protein 26..115 /note= "Mature PPSS-II' "
FT Misc-difference 74 /note= "Encoded by CAA"
FT Peptide 88..101 /note= "PPSS-II' pro-sequence"
FT Peptide 88..115 /note= "Prosomatostatin II' "
FT Cleavage-site 100..101 /note= "Dibasic cleavage site"
FT Peptide 102..115 /note= "SS-14 variant peptide"
XX
PN CA2325169-A1.
XX
D 03-JUN-2001.
XX
PF 01-DEC-2000; 2000CA-2325169.
XX
PR 03-DEC-1999; 99US-0168934.
XX
PA (NDSU-) NDSU RES FOUND.
XX
PI Sheridan MA, Moore CA, Kittelson JD;
XX
XX WPI; 2001-425997/46.
DR N-PSDB; AAS12934.
XX
PT New somatostatin polypeptides derived from Oncorhynchus mykiss, useful
PT for treating diabetes mellitus, acromegaly, gastrinoma, acquired
PT immunodeficiency syndrome and neurological disorders -
XX
PS Claim 2; Fig 3; 52pp; English.
XX
CC The invention relates to an Oncorhynchus mykiss somatostatin polypeptide
CC containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of
CC preprosomatostatin II (PPSS-II). The protein sequences and their
CC associated polynucleotides are useful for identifying modified
CC somatostatin polypeptide which functions as a somatostatin agonist useful

CC for research, therapeutics or diagnostics, including medical and
CC veterinary applications. The wild-type somatostatin and its modified
CC version are useful for treating hypersecretion from endocrine tumours in
CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.
CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage
CC through their effects on cell proliferation and apoptosis and as adjuncts
CC in the treatment of diabetes mellitus via inhibition of growth hormone
CC and glucagon. In addition, dysfunctional somatostatin secretion is
CC associated with acquired immunodeficiency syndrome (AIDS) and various
CC neurological disorders (e.g. epilepsy, Alzheimer's disease and
CC Huntington's disease) and somatostatin antagonists are effective in the
CC treatment of such conditions. Nucleic acids encoding the polypeptides are
CC useful in gene therapy and fusion peptides can be targeted to neoplasms
CC and their metastases, inhibiting the release of their secretory products.
CC This sequence represents O. Mykiss PPSS-II' protein.
CC Note: The features for this sequence are specifically claimed in the
CC specification.
XX
SQ Sequence 115 AA;
Query Match 77.6%; Score 45; DB 22; Length 115;
Best Local Similarity 88.9%; Pred. No. 1.1;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 DNLPPPERK 11
Db 93 nnlppperk 101
RESULT 7
ID AAM38940 standard; Protein; 290 AA.
XX
AC AAM38940;
XX 22-OCT-2001 (first entry)
XX Human polypeptide SEQ ID NO 2085.
DE Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR N-PSDB; AAI58096.
XX

PT Novel nucleic acids and polypeptides, useful for treating disorders
 XX such as central nervous system injuries -
 PS Example 3; SEQ ID NO 2085; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 290 AA;

Query Match 69.0%; Score 40; DB 22; Length 290;
 Best Local Similarity 72.7%; Pred. No. 24;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 11
 | | | | |
 Db 193 swdslppserk 203

RESULT 8
 ABB63556
 ID ABB63556 standard; Protein; 1006 AA.
 XX
 AC ABB63556;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 17460.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 JS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL07659.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 17460; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1006 AA;

Query Match 69.0%; Score 40; DB 22; Length 1006;
 Best Local Similarity 60.0%; Pred. No. 87;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 10
 | | | | |
 Db 444 atdmppprk 453

RESULT 9
 AAB07849
 ID AAB07849 standard; Protein; 219 AA.
 XX
 AC AAB07849;
 XX
 DT 14-NOV-2000 (first entry)

DE Amino acid sequence of a soybean type III glutathione-S-transferase.
 XX
 KW Soybean; glutathione-S-transferase; GST; detoxification;
 KW xenobiotic compound; herbicide-tolerance; transgenic plant;
 KW herbicide synergist.
 XX
 OS Glycine max.

PN WO200047728-A2.
 XX
 PD 17-AUG-2000.

PF 10-FEB-2000; 2000WO-US03347.
 XX
 XX 10-FEB-1999; 99US-0247373.

PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX McGonigle B, O'Keefe DP;

DR WPI; 2000-549144/50.
 DR N-PSDB; AAA59484.
 XX

PT Soybean glutathione-S-transferase polypeptides and polynucleotides used
 PT to produce herbicide tolerant transgenic plants and to screen for
 PT inhibitors or substrates of the enzyme -
 XX
 PS Claim 26; Page 81-82; 84pp; English.

XX The present sequence represents a soybean glutathione-S-transferase
 CC (GST) enzyme. The enzyme is involved in the detoxification of
 CC xenobiotic compounds in plants and seeds. The GST polynucleotides
 CC and polypeptides are used for the production of herbicide-tolerant
 CC transgenic plants, and for the development of screening assays to
 CC identify GST inhibitors and substrates, which can be used as
 CC herbicide synergists. GST Gene specific probes can be used in gene
 CC identification methods. The recombinant GST enzymes can be used to
 CC produce enzyme specific antibodies which are used to detect the
 CC enzymes in situ in cells or in vitro in cell extracts.
 XX
 SQ Sequence 219 AA;

Query Match 67.2%; Score 39; DB 21; Length 219;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10
Db 194 dnlprrdk 201

RESULT 10

AAB07840
ID AAB07840 standard; Protein; 224 AA.

XX AAB07840;

AC AAB07840;

XX 14-NOV-2000 (first entry)

DE Amino acid sequence of a soybean type III glutathione-S-transferase.
XX Soybean; glutathione-S-transferase; GST; detoxification;
KW xenobiotic compound; herbicide-tolerance; transgenic plant;
W herbicide synergist.

XX Glycine max.

OS WO200047728-A2.

XX 17-AUG-2000.

PD 10-FEB-2000; 2000WO-US03347.

XX 10-FEB-1999; 99US-0247373.

PR (DUPO) DU PONT DE NEMOURS & CO E I.

XX McGonigle B, O'Keefe DP;

DR WPI; 2000-549144/50.

DR N-PSDB; AAA59475.

XX Soybean glutathione-S-transferase polypeptides and polynucleotides used
PT to produce herbicide tolerant transgenic plants and to screen for
PT inhibitors or substrates of the enzyme -

XX Claim 26; Page 71-72; 84pp; English.

XX The present sequence represents a soybean glutathione-S-transferase
(GST) enzyme. The enzyme is involved in the detoxification of
CC xenobiotic compounds in plants and seeds. The GST polynucleotides
CC and polypeptides are used for the production of herbicide-tolerant
transgenic plants, and for the development of screening assays to
CC identify GST inhibitors and substrates, which can be used as
CC herbicide synergists. GST Gene specific probes can be used in gene
CC identification methods. The recombinant GST enzymes can be used to
CC produce enzyme specific antibodies which are used to detect the
enzymes in situ in cells or in vitro in cell extracts.

XX Sequence 224 AA;

Query Match 67.2%; Score 39; DB 21; Length 224;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10
Db 202 dnlprrdk 209

RESULT 11

AAG82843
ID AAG82843 standard; Protein; 244 AA.

XX

AC AAG82843;

XX 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:2780.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis;
KW vaccination; endocarditis.

XX Staphylococcus epidermidis.

XX WO200134809-A2.

XX 17-MAY-2001.

PD 09-NOV-2000; 2000WO-US30782.

XX 09-NOV-1999; 99US-0164258.

PR (GLAX) GLAXO GROUP LTD.

XX Kimmerly WJ;

XX WPI; 2001-316495/33.

DR N-PSDB; AAH53693.

XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -

XX Claim 18; Page 724; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
(II), given in AAG81454 to AAG83120; from Staphylococcus epidermidis
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
polynucleotide sequences from the present invention. AAH5091 to
CC AAH5098 represent oligonucleotide sequences and primers which are used
in the exemplification of the present invention.

CC N.B. The present invention specifically claims all the polynucleotide
sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.

XX Sequence 244 AA;

Query Match 67.2%; Score 39; DB 22; Length 244;
Best Local Similarity 70.0%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPERR 10
Db 177 saknlppkr 186

RESULT 12

ABB52820
ID ABB52820 standard; Protein; 370 AA.

XX ABB52820;

XX 11-FEB-2002 (first entry)

XX Escherichia coli polypeptide SEQ ID NO 1044.

XX KW Escherichia coli; B2/D+A-; antiinflammatory; antibacterial; immunosuppressive; extra-intestinal infection; phylogeny; meningitis; KW KW systemic infection; non-diarrhoeal infection; septicaemia; KW KW pyelonephritis; antibiotic resistance. XX OS Escherichia coli. XX XX WO200166572-A2. XX PN 13-SEP-2001. XX PD 12-MAR-2001; 2001WO-EP03445. XX PF 10-MAR-2000; 2000FR-0003145. XX PR 02-FEB-2001; 2001FR-0001449. XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE. XX I Bingen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C; XX DR WPI; 2001-550253/61. XX XX A library of DNA fragments of Escherichia coli strains for the PT PT phylogenetic determination of a given strain comprises polynucleotides of PT nature B2/D+ A- - XX XX Example 6; Fig 6; 646pp; English. XX PS The invention relates to a library of DNA fragments of Escherichia coli strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533) CC CC and encoded proteins (AB52459-AB52919 and AB52954-AB53094) of nature CC B2/D+A-. The polynucleotides have potential antiinflammatory, CC antibacterial and immunosuppressive activity as part of pharmaceutical CC compositions used to treat, palliate or prevent extra-intestinal E. coli CC infections. The polypeptides are useful for determining the phylogenic CC group of a given E. coli strain. These polypeptides can detect and treat CC an undesired development of E. coli, particularly an extra-intestinal CC infection that include systemic and non-diarrhoeal infections such as CC septicaemia, pyelonephritis and meningitis this is particularly CC advantageous as bacterial resistance is increasing with the more CC frequent use of broad spectrum antibiotics. XX SQ Sequence 370 AA;

Query Match 67.2%; Score 39; DB 22; Length 370; Best Local Similarity 100.0%; Pred. No. 47; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 VDNLPPR 8
Db 169 vdnlppr 175

RESULT 13
ABB69479
ID ABB69479 standard; Protein; 796 AA.
XX AC ABB69479;
XX XX 26-MAR-2002 (first entry).
XX DE Drosophila melanogaster polypeptide SEQ ID NO 35229.
XX KW Drosophila; developmental biology; cell signalling; insecticide; KW pharmaceutical.
XX OS Drosophila melanogaster.
XX XX WO200171042-A2.
XX PN 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.
XX XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX XX (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
DR N-PSDB; ABL13582.
XX XX New isolated nucleic acid detection reagent for detecting 1000 or more PT genes from Drosophila and for elucidating cell signalling and cell-cell PT interactions -
XX PS Disclosure; SEQ ID NO 35229; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent CC capable of detecting 1000 or more genes from Drosophila. The invention is CC useful in developmental biology and in elucidating cell signalling and CC cell-cell interactions in higher eukaryotes for the development of CC insecticides, therapeutics and pharmaceutical drugs. The invention CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA CC sequences (ABL01840-ABL16175) and the encoded proteins CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic format directly from WIPO CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 796 AA;

Query Match 67.2%; Score 39; DB 22; Length 796; Best Local Similarity 66.7%; Pred. No. 1e+02; Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 DNLPPRERK 11
Db 786 nnlpprgr 794

RESULT 14
AAW79680
ID AAW79680 standard; Protein; 33 AA.
XX AC AAW79680;
XX XX 11-JAN-1999 (first entry)
DT Synthetic CS198 derived peptide #4.
DE Gastrointestinal tract; GI tract; cancer; disease; detection; CS198; XX KW human; predisposition; treatment; Barret's oesophagus; gastric ulcer; KW gastritis; leiomyoma; polyps; Crohn's disease; ulcerative colitis; KW pancreatitis.
XX OS Synthetic.
XX XX WO9844159-A1.
XX PN 08-OCT-1998.
PD 30-MAR-1998; 98WO-US06251.
XX PF 31-MAR-1997; 97US-0828855.
XX PA (ABBO) ABBOTT LAB.
XX PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;
PI Gordon J, Granados EN, Hayden M, Hodges SC, Klass MR;
PI Kratochvil JD, Roberts-Rapp L, Russell JC, Stroupe SD;

XX DR WPI; 1998-542714/46.
XX CC
PT New gastrointestinal polynucleotides, CS198, and their detection -
PT used for developing products for the diagnosis and treatment of
PT gastrointestinal disorders, e.g. cancers, gastric ulcer or gastritis
XX PS
XX Claim 26; Page 103; 127pp; English.
XX CC
CC AAW79677-W79680 are synthetic CS198 derived peptide fragments which are
CC used in a method to detect the presence of a target human CS198
CC polynucleotide in a test sample. The CS198 gene is useful as a marker
CC for gastrointestinal (GI) tract disorders. The methods and products can
CC be used in detecting, diagnosing, staging, monitoring, prognosticating,
CC preventing or treating, or determining the predisposition to diseases
CC and conditions of the GI tract, such as GI tract cancer, Barrett's
CC oesophagus, gastric ulcer, gastritis, leiomyoma, polyps, Crohn's disease,
CC ulcerative colitis, and pancreatitis.
XX SQ Sequence 33 AA;

Query Match 65.5%; Score 38; DB 19; Length 33;
Best Local Similarity 63.6%; Pred. No. 5.7;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 11
|||: ||: ||
Db 3 svdsappgqrk 13

RESULT 15
AAE07323
ID AAE07323 standard; peptide; 33 AA.

AC AAE07323;

DT 06-NOV-2001 (first entry)

XX Human CS 198 peptide #4.

XX CS 198; gastrointestinal tract disease; GI tract; cancer; gastric ulcer;
KW gastritis; Crohn's disease; ulcerative colitis; pancreatitis;
KW Barrett's oesophagus; gene therapy; drug screening; human.

XX Homo sapiens.

XX US2001010904-A1.

XX 02-AUG-2001.

XX 30-MAR-1998; 98US-0050516.

XX 31-MAR-1997; 97US-0828855.

XX (BILL/) BILLING-MEDEL P A.

XX (COHE/) COHEN M.

XX (COLP/) COLPITTS T L.

XX (FRIE/) FRIEDMAN P N.

XX (GORD/) GORDON J.

XX (GRAN/) GRANADOS E N.

XX (HAYD/) HAYDEN M.

XX (HODG/) HODGES S C.

XX (KLAS/) KLASS M R.

XX (KRAT/) KRATOCHVIL J D.

XX (ROBE/) ROBERTS-RAPP L.

XX (RUSS/) RUSSELL J C.

XX (STRO/) STROUPE S D.

XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;

PI Granados EN, Hayden M, Hodges SC, Klass MR, Kratochvil JD;

PI Roberts-Rapp L, Russell JC, Stroupe SD;

XX

DR WPI; 2001-496163/54.

XX
PT Detecting the presence of target CS 198 polynucleotide, useful for
PT detecting or diagnosing diseases of the gastrointestinal tract,
PT comprises contacting test sample with at least one CS 198-specific
PT polynucleotide -
XX
XX Claim 17; Page 52; 68pp; English.

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Sequence 33 AA;

Query Match 65.5%; Score 38; DB 22; Length 33;

Best Local Similarity 63.6%; Pred. No. 5.7;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 11

|||: ||: ||

Db 3 svdsappgqrk 13

Search completed: June 13, 2002, 12:22:22

Job time: 275 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:17:47 ; Search time 72.45 Seconds
(without alignments)
174.775 Million cell updates/sec

Title: US-09-727-739B-3
Perfect score: 566
Sequence: 1 MLSTRVQCALALLSLALAI.....APRRKAGCKNFFWKFTSC.114

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	566	100.0	114	AAU07666	Rainbow trout prep
2	360	63.6	116	AAG03774	Human secreted pro
3	251	44.3	121	AAP20028	Sequence of prepro
4	177.5	31.4	115	AAU07667	Rainbow trout prep
5	164.5	29.1	111	AAU07668	Rainbow trout prep
6	162.5	28.7	125	AAP20029	Sequence of prepro
7	127	22.4	25	AAP20198	Sequence of somato
8	127	22.4	25	AAB91017	Somatostatin relat
9	127	22.4	28	AAW51859	Somatostatin analo
10	127	22.4	28	AAW28703	Mouse somatostatin
11	127	22.4	28	AAW24384	Somatostatin pepti

12	127	22.4	28	20	AAW24239	Peptide hormone so
13	127	22.4	28	22	AAB91018	Somatostatin relat
14	127	22.4	28	22	AAU07669	Mammalian somatost
15	127	22.4	29	22	AAB91020	Somatostatin relat
16	127	22.4	140	21	AAW69789	WWPsp-MWPmp20-(His
17	126	22.3	28	19	AAW50950	Somatostatin analo
18	126	22.3	28	22	AAB91021	Somatostatin relat
19	123	21.7	28	19	AAW50816	Somatostatin-28 an
20	122	21.6	28	3	AAP20130	Somatostatin-28 pe
21	122	21.6	28	3	AAP20131	Somatostatin-28 pe
22	120	21.2	28	3	AAP20197	Sequence of [D-Trp
23	119	21.0	28	19	AAW50818	Somatostatin-28 an
24	119	21.0	28	19	AAW50819	Somatostatin-28 an
25	119	21.0	28	19	AAW51855	Somatostatin analo
26	118	20.8	20	6	AAP50414	Swine duodenum eic
27	118	20.8	27	22	AAB91007	Somatostatin relat
28	118	20.8	28	4	AAP30058	Radioactively tagg
29	117	20.7	28	3	AAP20125	Somatostatin-28 pe
30	117	20.7	28	19	AAW51688	Somatostatin analo
31	113	20.0	28	3	AAP20126	Somatostatin-28 pe
32	111	19.6	28	3	AAP20128	Somatostatin-28 pe
33	111	19.6	28	19	AAW50817	Somatostatin-28 an
34	101	17.8	28	7	AAP61714	Somatostatin-28 an
35	101	17.8	28	10	AAP90989	Analogue of angler
36	101	17.8	32	22	AAB91009	Somatostatin relat
37	100	17.7	28	3	AAP20127	Somatostatin-28 pe
38	95	16.8	943	14	AAW34546	Somatostatin-leuko
39	94.5	16.7	112	19	AAW42033	Rat preprocartista
40	93.5	16.5	175	10	AAP90494	Dihydrofolate redu
41	93.5	16.5	175	10	AAP90496	Dihydrofolate redu
42	93	16.4	36	21	AAW58743	Somatostatin antig
43	93	16.4	54	21	AAW58760	Somatostatin antig
44	89.5	15.8	36	21	AAW58751	Somatostatin antig
45	89	15.7	24	20	AAW39551	CTLA-4 VLD CDRI re

ALIGNMENTS

RESULT 1
AAU07666
ID AAU07666 standard; Protein: 114 AA.
XX
AC AAU07666;
XX
DT 04-DEC-2001 (first entry)
XX
DE Rainbow trout preprosomatostatin I (PPSS-I) polypeptide.
XX
KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
KW PPSS-II'; PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;
KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;
KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;
KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.
OS Oncorhynchus mykiss.

Key	Location/Qualifiers
Peptide	1..24
Protein	/note= "Signal peptide"
Protein	1..88
Protein	/note= "PPSS-I pre-sequence"
Protein	25..114
Protein	/note= "Mature PPSS-I"
Protein	89..100
Protein	/note= "PPSS-I pro-sequence"
Protein	89..114
Protein	/note= "Prosomatostatin I"
Protein	99..100
Protein	/note= "Dibasic cleavage site"

XX PD 03-MAR-1982. 44.3%; Score 251; DB 3; Length 121;
XX XX 21-AUG-1981; 81EP-0303825. 51.3%; Pred. No. 3.6e-20;
PF XX 25-AUG-1980; 80US-0181046. 15; Mismatches 36; Indels 6; Gaps 4;
XX XX (REGC) UNIV OF CALIFORNIA.
XX PI Hobart P, Crawford R, Pictet RL, Rutter WJ;
XX WPI; 1982-18113E/10.
DR N-PSDB; AAN20033.
XX New somatostatin and precursors - produced by transformed
PT microorganisms
XX Example; Fig 3; 50pp; English.
XX The inventors claim preprosomatostatin-1, prosomatostatin-1,
XX preprosomatostatin-2, prosomatostatin-2 and somatostatin-2; and DNA
XX encoding them. The translation of somatostatin mRNA yields a
XX precursor (prepro S1) containing a signal peptide which may be
XX released during the transit into the endoplasmic reticulum, and the
XX resultant precursor (pro S1) is subsequently cleaved to yield S1
XX itself. The prepeptide portion of prepro S1 is probably about 20-25
XX bases long. Translation of plaS2 predicts the sequence of a 125 AA
XX peptide which surprisingly contains a 14 AA sequence at its carboxy
XX terminus which differs from S1 by only 2 AAs, and is termed
XX Somatostatin 2 (S2).
XX Sequence 121 AA;
SQ

Query Match 44.3%; Score 251; DB 3; Length 121;
Best Local Similarity 51.3%; Pred. No. 3.6e-20;
Matches 60; Conservative 15; Mismatches 36; Indels 6; Gaps 4;
QY 3 STRVOCALA-LLSLALAISSVSAPSDAKLRLQLQSLMAPAGKQLARNTLVE-LLSEL 60
Db 6 ssrlcllvllsltasiscsfaggrdskrlrlhrypl-qgskqdmtrsalaeillsdl 64
QY 61 AHVENEAEIELDDM---SHGVEQEDVDLELERAPGPVPLAPRERKAGCKNFFWKTFTSC 114
Db 65 lggenealeeniflaeggpedahadleraaaggppllaprerkgagknffwktftsc 121

--RESULT 4
AU07667
D AU07667 standard; Protein; 115 AA.
XX AU07667;
AC AU07667;
XX AU07667;
DT 04-DEC-2001 (first entry)
XX Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.
DE
XX Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
KW PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;
KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;
KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;
KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.
XX
OS Oncorhynchus mykiss.
XX
XX Key Location/Qualifiers
FH Peptide 1..25
FT /note= "Signal peptide"
FT 1..87
FT Protein /note= "PPSS-II' pre-sequence"

FT Protein 26..115 /note= "Mature PPSS-II'"
FT Misc-difference 74 /note= "Encoded by CAA"
FT Peptide 88..101 /note= "PPSS-II' pro-sequence"
FT Peptide 88..115 /note= "Prosomatostatin II'"
FT Cleavage-site 100..101 /note= "Dibasic cleavage site"
FT Peptide 102..115 /note= "SS-14 variant peptide"
XX CA2325169-A1.
XX 03-JUN-2001.
XX 01-DEC-2000; 2000CA-2325169.
XX 03-DEC-1999; 99US-0168934.
XX (NDSU-) NDSU RES FOUND.
XX Sheridan MA, Moore CA; Kittelson JD;
XX WPI; 2001-425997/46.
XX N-PSDB; AAS12934.
XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful
XX for treating diabetes mellitus, acromegaly, gastrinoma, acquired
XX immunodeficiency syndrome and neurological disorders -
XX Claim 2; Fig 3; 52pp; English.
XX The invention relates to an Oncorhynchus mykiss somatostatin polypeptide
XX containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of
XX preprosomatostatin II (PPSS-II). The protein sequences and their
XX associated polynucleotides are useful for identifying modified
XX somatostatin polypeptides which functions as a somatostatin agonist useful
XX for research, therapeutics or diagnostics, including medical and
XX veterinary applications. The wild-type somatostatin and its modified
XX version are useful for treating hypersecretion from endocrine tumours in
XX the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.
XX gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage
XX through their effects on cell proliferation and apoptosis and as adjuncts
XX in the treatment of diabetes mellitus via inhibition of growth hormone
XX and glucagon. In addition, dysfunctional somatostatin secretion is
XX associated with acquired immunodeficiency syndrome (AIDS) and various
XX neurological disorders (e.g. epilepsy, Alzheimer's disease and
XX Huntington's disease) and somatostatin antagonists are effective in the
XX treatment of such conditions. Nucleic acids encoding the polypeptides are
XX useful in gene therapy and fusion peptides can be targeted to neoplasms
XX and their metastases, inhibiting the release of their secretory products.
XX This sequence represents O. Mykiss PPSS-II' protein.
XX Note: The features for this sequence are specifically claimed in the
XX specification.
XX Sequence 115 AA;
SQ

Query Match 31.4%; Score 177.5; DB 22; Length 115;
Best Local Similarity 34.7%; Pred. No. 5.2e-12;
Matches 50; Conservative 9; Mismatches 18; Indels 67; Gaps 4;
QY 5 RVQCALALLSLALAISSVSAPSS---DAKLRLQLQLRSLMA----- 41
Db 5 rihcalallglalaicsggaasqgdldlrslrllqraaaalphrsgvserwrtfypncp 64
QY 42 -----PAGKQELARNTLVLELSELAHVENEAEIELDDMSHGVEQEDVDLELERAPGP 92
Db 65 clrprkvkcpag-----akdlrvelersvgn 91
QY 93 V--LAPRERKAGCKNFFWKTFTSC 114

Db 92 pnnlpprerkagcknfywkgftsc 115
RESULT 5
AAU07668
ID AAU07668 standard; Protein; 111 AA.
XX
AC AAU07668;
XX
DT 04-DEC-2001 (first entry)
XX
DE Rainbow trout preprosomatostatin II (PPSS-II'') polypeptide.
XX
KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
KW PPSS-II'; PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;
KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;
KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;
KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.
XX
OS Oncorhynchus mykiss.
XX
FH Key Location/Qualifiers
FT Peptide 1..25
FT /note= "Signal peptide"
FT Protein 1..86
FT /note= "PPSS-II'' pre-sequence"
FT Protein 26..111
FT /note= "Mature PPSS-II''"
FT Peptide 87..97
FT /note= "PPSS-II'' pro-sequence"
FT Peptide 87..111
FT /note= "Prosomatostatin II''"
FT Cleavage-site 96..97
FT /note= "Dibasic cleavage site"
FT Peptide 98..111
FT /note= "SS-14 variant peptide"
XX
CA2325169-A1.
XX
PD 03-JUN-2001.
XX
PF 01-DEC-2000; 2000CA-2325169.
XX
QR 03-DEC-1999; 99US-0168934.
XX
A (NDSU-) NDSU RES FOUND.
XX
PI Sheridan MA, Moore CA, Kittelson JD;
XX
DR WPI; 2001-425997/46.
DR N-PSDB; AAS12935.
XX
PT New somatostatin polypeptides derived from Oncorhynchus mykiss, useful
PT for treating diabetes mellitus, acromegaly, gastrinoma, acquired
PT immunodeficiency syndrome and neurological disorders -
XX
PS Claim 1; Fig 3; 52pp; English.
XX
CC The invention relates to an Oncorhynchus mykiss somatostatin polypeptide
CC containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of
CC preprosomatostatin II (PPSS-II). The protein sequences and their
CC associated polynucleotides are useful for identifying modified
CC somatostatin polypeptide which functions as a somatostatin agonist useful
CC for research, therapeutics or diagnostics, including medical and
CC veterinary applications. The wild-type somatostatin and its modified
CC version are useful for treating hypersecretion from endocrine tumours in
CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.
CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage
CC through their effects on cell proliferation and apoptosis and as adjuncts

CC in the treatment of diabetes mellitus via inhibition of growth hormone
CC and glucagon. In addition, dysfunctional somatostatin secretion is
CC associated with acquired immunodeficiency syndrome (AIDS) and various
CC neurological disorders (e.g. epilepsy, Alzheimer's disease and
CC Huntington's disease) and somatostatin antagonists are effective in the
CC treatment of such conditions. Nucleic acids encoding the polypeptides are
CC useful in gene therapy and fusion peptides can be targeted to neoplasms
CC and their metastases, inhibiting the release of their secretory products.
CC This sequence represents O. Mykiss PPSS-II'' protein.
CC Note: The features for this sequence are specifically claimed in the
CC specification.
XX
SQ Sequence 111 AA;
Query Match 29.1%; Score 164.5; DB 22; Length 111;
Best Local Similarity 38.7%; Pred. No. 1.4e-10;
Matches 48; Conservative 15; Mismatches 38; Indels 23; Gaps 5;
QY 1 MLSTRVQCALALLSLALAISSVSAAPS---DAKLRLQLRSLMAP---AGKQELARNTL 53
Db 1 mrvsqihcalallglalaicsgaasqpdldasrrllgralaalphrsgvserwrtfy 60
QY 54 VE---LLSELAHVENEAIELDDMSHGEVDLELERAPGVLAPRERKAGCKNFFWKT 110
Db 61 pncpclrwrprkvkpgqlk-----akedlersvdlp-----prerkagcknfywkg 107
QY 111 FTSC 114
Db 108 ftsc 111
RESULT 6
AAP20029
ID AAP20029 standard; Protein; 125 AA.
XX
AC AAP20029;
XX
DT 14-AUG-1992 (first entry)
XX
DE Sequence of preprosomatostatin-2 encoded on plas2.
XX
KW Somatostatin; growth hormone; peptide hormone; secretion.
XX
FH Key Location/Qualifiers
FT Protein 112..125
FT /label= Somatostatin II
XX
PN EP46669-A.
XX
PD 03-MAR-1982.
XX
PF 21-AUG-1981; 81EP-0303825.
XX
PR 25-AUG-1980; 80US-0181046.
XX
PA (REGC) UNIV OF CALIFORNIA.
XX
PI Hobart P, Crawford R, Pictet RL, Rutter WJ;
XX
DR WPI; 1982-18113E/10.
DR N-PSDB; AAN20034.
XX
PT New somatostatin and precursors - produced by transformed
PT microorganisms
XX
PS Example; Fig 3; 50pp; English.
XX
CC The inventors claim preprosomatostatin-1, prosomatostatin-1,
CC preprosomatostatin-2, prosomatostatin-2 and somatostatin-2; and DNA
CC encoding them. The translation of somatostatin mRNA yields a
CC precursor (prepro S1) containing a signal peptide which may be
CC released during the transit into the endoplasmic reticulum, and the

CC resultant precursor (pro S1) is subsequently cleaved to yield S1
CC itself. The prepeptide portion of prepro S1 is probably about 20-25
CC bases long. Translation of plas2 predicts the sequence of a 125 AA
CC peptide which surprisingly contains a 14 AA sequence at its carboxy
CC terminus which differs from S1 by only 2 AAs, and is termed
CC Somatostatin 2 (S2).

XX Sequence 125 AA;

Query Match 28.7%; Score 162.5; DB 3; Length 125;
Best Local Similarity 38.9%; Pred. No. 2.7e-10;
Matches 51; Conservative 16; Mismatches 41; Indels 23; Gaps 6;

QY 1 MLSTRVQCALALLSLAIAISSVSA-----APSDAKLRQ--LLQRLMAPAGKQELA 49
Db 1 mqclrcpailallalvlgcpvssqldreqsdnqldlelrqhwllersagsllsqews 60
QY 50 RNTVELLSEL----AHVENEAIELDDMSHGVEQEDVDLELERAPGPV--LAPRERKAGC 103
b 61 kraveellaqmslpeatfqrea---edasmategr---mnlersvdstnnlpprerkgac 114
QY 104 KNFFWKFTTSC 114
Db 115 knfywkgftsc 125

RESULT 7

AAP20198
ID AAP20198 standard; Protein; 25 AA.

XX
AC AAP20198;

XX 14-AUG-1992 (first entry)

XX Sequence of somatostatin-25 analogue.

XX Somatostatin; hormone; growth hormone release; inhibition.

XX Key Location/Qualifiers

FT Modified-site 1

FT Disulfide-bond 14..25

FT Modified-site 25

FT /label= C-OH

XX US4316891-A.

XX 23-FEB-1982.

XX 14-JUN-1980; 80US-0159801.

XX 14-JUN-1980; 80US-0159801.

XX (SALK-) SALK INST BIOLOG.

XX Guillemin RCL, Esch FS, Bohlen P, Brazeau PE, Ling NC;

XX WPI; 1982-19801E/10.

XX Extended somatostatin analogues - with increased inhibition of

XX growth hormone release

XX Claim 6; Column 12; 7pp; English.

XX The inventors claim a pharmaceutical compsn. which comprises (i)
CC synthetic somatostatin-28 (SS-28), SS-25 or (D-Trp(22))-SS-28 and
CC (ii) a liq. or solid carrier, and SS-28 derivs. and SS-25 derivs.
CC The compsn. and derivs. are more potent than somatostatin in
CC inhibiting release of growth hormone; they also inhibit basal and
CC stimulated insulin and glucagon secretion. (D-Trp(22))-SS-28
CC exhibits very substantial increases in potency w.r.t. inhibition of
CC growth hormone secretion.

XX Sequence 25 AA;

Query Match 22.4%; Score 127; DB 3; Length 25;
Best Local Similarity 91.3%; Pred. No. 3e-07;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 92 PVLAPRERKAGCKNFFWKFTTSC 114
Db 3 pamaprerkagcknffwkftsc 25

RESULT 8

AAB91017

ID AAB91017 standard; Peptide; 25 AA.

XX AAB91017;

XX 22-JUN-2001 (first entry)

XX Somatostatin related peptide SEQ ID NO:191.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;
XX blood component; modification; succinimidyl; maleimido group; amino;
XX hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX Homo sapiens.

XX Synthetic.

XX WO200069900-A2.

XX 23-NOV-2000.

XX 17-MAY-2000; 2000WO-US13576.

XX 17-MAY-1999; 99US-0134406.

XX 10-SEP-1999; 99US-0153406.

XX 15-OCT-1999; 99US-0159783.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents
XX peptidase degradation, useful for increasing length of in vivo activity

XX Disclosure; Page 252; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.

XX Sequence 25 AA;

Query Match	22.4%	Score 127;	DB 22;	Length 25;
Best Local Similarity	91.3%	Pred. NO. 3e-07;		
Matches 21;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
QY	92	PVLAPRERKAGCKNFEFWKFTTSC	114	
		:		
Db	3	pamaprerkagcknfwkfttsc	25	

RESULT	9
AAW51859	
ID	AAW51859 standard; peptide; 28 AA.
XX	
AC	AAW51859;
XX	
DT	10-SEP-1998 (first entry)
XX	
DE	Somatostatin analogue for the treatment of syndrome X of Reaven.
XX	
W	Somatostatin analogue; syndrome X of Reaven; hyperinsulinaemia syndrome; diazoxide; cyclothiazide; metformin.

The invention relates to a pharmaceutical composition for treatment of the risk factors of syndrome X of Reaven (hyperinsulinaemia syndrome). It comprises somatostatin, diazoxide, cyclothiazide (or an analogue of one of these) or metformin as the active ingredient. The composition reduces resistance to insulin, and so treats and prevents all the associated risk factors at once. The risk factors are hypertension, dyslipidaemia (raised triglyceride and LDL levels with reduced HDL levels), shorter coagulation time due to increased Plasminogen Activator Inhibitor-1 levels, core obesity, glucose intolerance hyperinsulinaemia. The composition reduces the incidence of ischaemic heart disease, cerebrovascular disorders, intermittent claudication, ischaemic bowel disease, impotence due to peripheral vascular disease, hypercoagulation (e.g. renal vein thrombosis), obesity and glucose intolerance. The present sequence represents a specifically claimed somatostatin analogue.

RESULT 10

AA	Y28703	standard; peptide; 28 AA.
AA	Y28703	
AA	Y28703	
DT	07-OCT-1999	(first entry)
DE	Mouse somatostatin SS-28	hormone.
KW	Mouse somatostatin SS-28	hormone; growth hormone; insulin; glucagon;
KW	thyroid stimulating hormone;	octreotide; cell-based delivery of insulin;
KW	glucose-stimulated insulin	secretion; SSTRV; somatostatin receptor;
KW	mouse somatostatin receptor	type V gene; diabetes.
OS	Mus musculus.	
PN	WO9935242-A1.	
PD	15-JUL-1999.	
PF	11-JAN-1999;	99WO-US00633.
PR	03-JUN-1998;	98US-0087848.
PR	12-JAN-1998;	98US-0071193.
PR	12-JAN-1998;	98US-0071209.
PR	12-JAN-1998;	98US-0072556.
PR	03-JUN-1998;	98US-0087821.
PA	(BETA-) BETAGENE INC.	
PI	Clark SA, Quaade C;	
PI	WPI; 1999-444195/37.	
PT	New defined medium for	culture of neuroendocrine cells, e.g. of
PT	insulin-secreting cells	
PS	Example 8; Page 143;	312pp; English.
SS	The present sequence is	a mouse somatostatin (SS-28) hormone which
CC	was found to inhibit the	release of growth hormone, thyroid stimulating
CC	hormone, insulin and	glucagon. In addition, SS-28 and its analogue
CC	Octreotide may inhibit	growth of some tumours. The hormone was
CC	used to study its effect	on glucose-stimulated insulin secretion in
CC	high expressing and	non-expressing clones of mouse somatostatin
CC	receptor, type V gene	(SSTRV). The insulin secretion was highly inhibited
CC	in the high expressing	clone as compared to the non-expressing clone
CC	because the high	expressing clone showed high sensitivity to
CC	somatostatin. The	hormone effectively inhibits insulin secretion in the
CC	absence of glucose.	The somatostatin receptor can be introduced in cell
CC	lines, used in cell-	based delivery of insulin for treating diabetes, for
CC	precise regulation	of insulin release.
SS	Sequence	28 AA;

RESULT 11
 AAY24384
 ID AAY24384 standard; peptide; 28 AA
 XX
 AC AAY24384;
 XX
 DT 20-SEP-1999 (first entry)
 XX

DE XX Somatostatin peptide hormone SS-28.
KW XX Glucagon-like peptide I receptor; GLP-1 receptor; drug screening;
KW XX secretory function; immortalized neuroendocrine secretory cell;
KW XX regulation; diabetes; insulin secretion; neuroendocrine-based disorder;
KW XX Parkinson's disease; atypical cretinism; Addison's disease.
OS Mus musculus.
PN WO9935495-A2.
XX
PD 15-JUL-1999.
XX
PF 11-JAN-1999; 99WO-US00551.
XX
PR 03-JUN-1998; 98US-0087848.
PR 12-JAN-1998; 98US-0071193.
PR 12-JAN-1998; 98US-0071209.
PR 12-JAN-1998; 98US-0072556.
PR 03-JUN-1998; 98US-0087821.
XX
PA (BETA-) BETAGENE INC.
XX
PI Clark SA, Thigpen AE;
XX
DR WPI; 1999-430454/36.
XX
PT New modulators of secretory function, used to control peptide
PT secretion from cells in vivo or in vitro, specifically for treating
PT diabetes
XX
PS Example 8; Page 153-154; 309pp; English.
XX
XX The present invention describes a method for identifying modulators (I)
CC of secretory function by treating an immortalized cell, having a stable
CC secretory function, with a test compound and detecting any change in
CC secretion caused by the compound. (I) are used to control secretion of
CC polypeptides from cells, in vivo or in vitro. Specifically they are used
CC for treating or preventing diabetes by regulation of insulin secretion,
CC but can also be used in cases of other neuroendocrine-based disorders
CC such as Parkinson's disease, atypical cretinism and Addison's disease.
CC The method uses engineered, immortalized cells that are available in
CC large amounts, with a stable and predictable phenotype. They allow
CC screening to be performed in vivo. The present sequence represents a
CC somatostatin peptide hormone SS-28 used in an example from the present
CC invention.
XX
XX
XX

OS Synthetic.
XX WO9935255-A2.
PN
XX
PD 15-JUL-1999.
XX
PF 11-JAN-1999; 99WO-US00631.
XX
PR 03-JUN-1998; 98US-0087848.
PR 12-JAN-1998; 98US-0071193.
PR 12-JAN-1998; 98US-0071209.
PR 12-JAN-1998; 98US-0072556.
PR 03-JUN-1998; 98US-0087821.
XX
PA (BETA-) BETAGENE INC.
XX
PI Clark SA, Thigpen AE;
XX
DR WPI; 1999-419351/35.
XX
PT New immortalized neuroendocrine cells that stably secrete
PT polypeptide, particularly used to treat diabetes and hypoglycaemia
XX
PS Example 8; Page 148; 318pp; English.
XX
XX The present invention describes immortalized neuroendocrine cells (A)
CC that stably secrete a polypeptide hormone (I) contain an expression
CC region that includes a transgene (TG), linked to a promoter functional
CC in eukaryotic cells, such that expression of TG increases sensitivity
CC of the cells to a modulator of (I) secretion. (A) are specifically
CC used, by transplantation, to treat diabetes or hypoglycaemia (especially
CC where associated with insulin therapy) but more generally are used to
CC express, in vivo, a wide range of therapeutic hormones, enzymes,
CC amidated proteins and growth factors. Also engineered neuroendocrine
CC cells are used to identify new therapeutic agents or drug targets.
CC (A) have a stable phenotype and particularly inducible glucagon
CC secretion and glucose counter-regulatory capacities, i.e. they balance
CC the hyperglycaemic effects of beta-cell loss and the hypoglycaemic
CC effects of administered insulin. Since they are of human origin, they
CC are less likely to suffer immune rejection than xenografts. The present
CC sequence represents the peptide hormone somatostatin SS-28 used in an
CC example from the present invention.
XX
XX Sequence 28 AA;
SQ

Query Match 22.4%; Score 127; DB 20; Length 28;
Best Local Similarity 91.3%; Pred. No. 3.5e-07;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 92 PVLAPRRKAGCKNFFWKTFTSC 114
| :|||||
Db 6 pamaprerkagcknffwktftsc 28

RESULT 13
AAB91018
ID AAB91018 standard; peptide; 28 AA.
XX
AC AAB91018;
XX
DT 22-JUN-2001 (first entry)
XX
DE Somatostatin related peptide SEQ ID NO:192.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimide; maleimide group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.

Query Match 22.4%; Score 127; DB 20; Length 28;
Best Local Similarity 91.3%; Pred. No. 3.5e-07;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 92 PVLAPRRKAGCKNFFWKTFTSC 114
| :|||||
Db 6 pamaprerkagcknffwktftsc 28

RESULT 12
AAY24239
ID AAY24239 standard; peptide; 28 AA.
XX
AC AAY24239;
XX
DT 15-SEP-1999 (first entry)
XX
DE Peptide hormone somatostatin SS-28.

XX Glucagon-like peptide I receptor; GLP-1; regulation; secretion;
KW neuroendocrine cell line; glycaemic sensing mechanism; glucose;
KW genetic engineering; hypoglycaemia; diabetes.
XX

XX PD 23-NOV-2000.
XX XX
XX PF 17-MAY-2000; 2000WO-US13576.
XX XX
XX PR 17-MAY-1999; 99US-0134406.
XX PR 10-SEP-1999; 99US-0153406.
XX PR 15-OCT-1999; 99US-0159783.
XX PA (CONJ-) CONJUCHEM INC.
XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX WPI; 2001-112059/12.
XX PT Modifying and attaching therapeutic peptides to albumin prevents
XX PT peptidase degradation, useful for increasing length of in vivo activity
XX PT
XX
XX Disclosure; Page 252; 733pp; English.

CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimide and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.

XX Sequence 28 AA;

Query Match 22.4%; Score 127; DB 22; Length 28;
Best Local Similarity 91.3%; Pred. No. 3.5e-07;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 92 PVLAPRERKAGCKNFFWKTFTSC 114
I :|||||
6 pamaprerkagcknffwktftsc 28

RESULT 14
AAU07669
ID AAU07669 standard; Peptide; 28 AA.

XX AC AAU07669;

XX DT 04-DEC-2001 (first entry)

XX DE Mammalian somatostatin 28 (SS-28) peptide sequence.

XX KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
XX KW PPSS-II'; PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
XX KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
XX KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone; SS-28;
XX KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
XX KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;
XX KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;
XX KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.

XX OS Mammalia sp.

XX XX CA2325169-A1.

XX PD 03-JUN-2001.
XX XX
XX PF 01-DEC-2000; 2000CA-2325169.
XX XX
XX PR 03-DEC-1999; 99US-0168934.
XX XX
XX PA (NDSU-) NDSU RES FOUND.
XX PI Sheridan MA, Moore CA, Kittelson JD;
XX WPI; 2001-425997/46.
XX
XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful
XX PT for treating diabetes mellitus, acromegaly, gastrinoma, acquired
XX PT immunodeficiency syndrome and neurological disorders
XX
XX Example 5; Fig 6; 52pp; English.

CC The invention relates to an Oncorhynchus mykiss somatostatin polypeptide
CC containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of
CC preprosomatostatin II (PPSS-II). The protein sequences and their
CC associated polynucleotides are useful for identifying modified
CC somatostatin polypeptide which functions as a somatostatin agonist useful
CC for research, therapeutics or diagnostics, including medical and
CC veterinary applications. The wild-type somatostatin and its modified
CC version are useful for treating hypersecretion from endocrine tumours in
CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.
CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage
CC through their effects on cell proliferation and apoptosis and as adjuncts
CC in the treatment of diabetes mellitus via inhibition of growth hormone
CC and glucagon. In addition, dysfunctional somatostatin secretion is
CC associated with acquired immunodeficiency syndrome (AIDS) and various
CC neurological disorders (e.g. epilepsy, Alzheimer's disease and
CC Huntington's disease) and somatostatin antagonists are effective in the
CC treatment of such conditions. Nucleic acids encoding the polypeptides are
CC useful in gene therapy and fusion peptides can be targeted to neoplasms
CC and their metastases, inhibiting the release of their secretory products.
CC This sequence represents the mammalian somatostatin 28 (SS-28) peptide.

XX Sequence 28 AA;

Query Match 22.4%; Score 127; DB 22; Length 28;
Best Local Similarity 91.3%; Pred. No. 3.5e-07;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 92 PVLAPRERKAGCKNFFWKTFTSC 114
I :|||||
6 pamaprerkagcknffwktftsc 28

RESULT 15
AAB91020
ID AAB91020 standard; Peptide; 29 AA.

XX AC AAB91020;

XX DT 22-JUN-2001 (first entry)

XX DE Somatostatin related peptide SEQ ID NO:194.

XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
XX KW blood component; modification; succinimide; maleimido group; amino;
XX KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200069900-A2.

XX PD 23-NOV-2000.

XX XX

PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
PR 10-SEP-1999; 99US-0153406.
PR 15-OCT-1999; 99US-0159783.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT
XX
PS Disclosure; Page 253; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 29 AA;

Query Match 22.4%; Score 127; DB 22; Length 29;
Best Local Similarity 91.3%; Pred. No. 3.7e-07;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 92 PVLAPRERKAGCKNFFWKFTSC 114
| :|||||
Db 7 pamaprerkagcknffwktftsc 29

Search completed: June 13, 2002, 12:22:20
Job time: 273 sec

QY 67 AIELDDMSHGVEQEDVDLELERA--PGPVLAPRRKAGCKNFFWKFTTSC 114
Db 61 ALEPEDLPQAAEQDEMRLQSRANSNPAMAPRRKAGCKNFFWKFTTSC 110

RESULT 2
US-08-648-322-2
; Sequence 2, Application US/08648322
; Patent No. 6074872
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, Gregor J.
; APPLICANT: de lecea, Luis
; TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
; STREET: 10666 No. 6074872th Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/648,322
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: 519.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 554-2937
; TELEFAX: (619) 554-6312
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-648-322-2

Query Match 16.7%; Score 94.5; DB 3; Length 112;
; Best Local Similarity 33.0%; Pred. No. 0.00055;
Matches 37; Conservative 13; Mismatches 43; Indels 19; Gaps 6;
QY 9 ALALLSLALAISVV--SAAPSDAKLRLQLQSLMAPAGK---QELARNTLVLLSELAHV 63
Db 13 ALSLL-LLLLSGIAASALPLES-----GPTGQDSVQDATGGRRTGLTFLAWW 60
QY 64 ENEAIELDDMSHGVEQEDVDLELERAPGPVLAPRR-KAGCKNFFWKFTTSC 114
Db 61 HEWASQ-DSSSTAPEGGTPELSKROERPPLOQPPHRRDKKPKCNFFWKFTFSSC 111

RESULT 3
US-09-001-472-3
; Sequence 3, Application US/09001472
; Patent No. 6232100
; GENERAL INFORMATION:
; APPLICANT: OLSEN, HENRIK S.
; APPLICANT: RUBEN, STEVEN M.
; TITLE OF INVENTION: CORTISTATIN POLYPEPTIDES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600

; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/001,472
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,980
; FILING DATE: 31-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/037,386
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0430002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-001-472-3

Query Match 16.7%; Score 94.5; DB 4; Length 112;
; Best Local Similarity 33.0%; Pred. No. 0.00055;
Matches 37; Conservative 13; Mismatches 43; Indels 19; Gaps 6;
QY 9 ALALLSLALAISVV--SAAPSDAKLRLQLQSLMAPAGK---QELARNTLVLLSELAHV 63
Db 13 ALSLL-LLLLSGIAASALPLES-----GPTGQDSVQDATGGRRTGLTFLAWW 60
QY 64 ENEAIELDDMSHGVEQEDVDLELERAPGPVLAPRR-KAGCKNFFWKFTTSC 114
Db 61 HEWASQ-DSSSTAPEGGTPELSKROERPPLOQPPHRRDKKPKCNFFWKFTFSSC 111

RESULT 4
US-08-455-970A-10
; Sequence 10, Application US/08455970A
; Patent No. 5708155
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,970A

Db 1 AGCKNFFWKTFTSC 14

RESULT 7

US-08-416-007-4
Sequence 4, Application US/08416007
Patent No. 5693679
GENERAL INFORMATION:
APPLICANT: Vincent, Jean-Pierre
APPLICANT: Gaudriault, Georges
APPLICANT: Beaudet, Alain
TITLE OF INVENTION: FLUORESCENT SOMATOSTATIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416.007
FILING DATE: 04-APR-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06942/003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-416-007-4

Query Match 15.4%; Score 87; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKTFTSC 114
Db 1 AGCKNFFWKTFTSC 14

RESULT 8

US-08-676-263-11
Sequence 11, Application US/08676263
Patent No. 5705143
GENERAL INFORMATION:
APPLICANT: Bower, Gary R.
APPLICANT: Forster, Alan M.
APPLICANT: Riley, Anthony E.
APPLICANT: Storey, Anthony E.
TITLE OF INVENTION: BIOLOGICAL TARGETING AGENTS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America

ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/676.263
FILING DATE: 07-NOV-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94300224.6
FILING DATE: 12-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sharp, Jeffrey S.
REGISTRATION NUMBER: 31,879
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: (312) 474-6600
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Cross-links
LOCATION: 3.14
US-08-676-263-11

Query Match 15.4%; Score 87; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKTFTSC 114
Db 1 AGCKNFFWKTFTSC 14

RESULT 9

US-08-286-748B-13
Sequence 13, Application US/08286748B
Patent No. 5759542
GENERAL INFORMATION:
APPLICANT: Victor Gurewich
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY OF DRUGS BY PLATELETS FOR THE TREATMENT OF CARDIOVASCULAR AND OTHER DISEASES
TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF CARDIOVASCULAR AND OTHER DISEASES
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286.748B
FILING DATE: August 5, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: J. Peter Fasse

REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04547/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 14
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-286-748B-13

Query Match 15.4%; Score 87; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKFTTSC 114
|||||
Db 1 AGCKNFFWKFTTSC 14

RESULT 10
US-08-690-090A-1
Sequence 1, Application US/08690090A.
Patent No. 5770687
GENERAL INFORMATION:
APPLICANT: HORNIK, VERED
APPLICANT: SERI-LEVY, ALON
APPLICANT: GELLERMAN, GARY
APPLICANT: GILON, CHAIM
TITLE OF INVENTION: Conformationally Constrained Backbone
TITLE OF INVENTION: Cyclized Somatostatin Analogs
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,090A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/488,159
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7754-052-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-690-090A-1

Query Match 15.4%; Score 87; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 AGCKNFFWKFTTSC 114
|||||
Db 1 AGCKNFFWKFTTSC 14
RESULT 11
US-08-488-159-1
Sequence 1, Application US/08488159
Patent No. 5811392
GENERAL INFORMATION:
APPLICANT: Gilon, Chaim
TITLE OF INVENTION: Conformationally Constrained Backbone
TITLE OF INVENTION: Cyclized Peptide Analogs
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,159
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7754-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-159-1

Query Match 15.4%; Score 87; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKFTTSC 114
|||||
Db 1 AGCKNFFWKFTTSC 14

RESULT 12
US-08-465-764-1
Sequence 1, Application US/08465764
Patent No. 5814298
GENERAL INFORMATION:
APPLICANT: Dean, Richard T
APPLICANT: Lister-James, John
TITLE OF INVENTION: Technetium-99m Labeled
TITLE OF INVENTION: Somatostatin-derived Peptides for Imaging and Therapeutic
TITLE OF INVENTION: Uses
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Banner & Allegretti, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,764
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5814298nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 92,385-R
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3..14
; OTHER INFORMATION: /label= Disulfide bond
; OTHER INFORMATION: /note= "The peptide is cyclized between the
; OTHER INFORMATION: sidechain sulfur atoms of the 3d and 14th residues
US-08-465-764-1

Query Match 15.4%; Score 87; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKFTTSC 114
| | | | | | | | | | | | | | | |
Db 1 AGCKNFFWKFTTSC 14

-----RESULT 13
; -08-475-751-4
; Sequence 4, Application US/08475751
; Patent No. 5824772
; GENERAL INFORMATION:
; APPLICANT: Vincent, Jean-Pierre
; APPLICANT: Gaudriault, Georges
; APPLICANT: Beaudet, Alain
; TITLE OF INVENTION: FLUORESCENT SOMATOSTATIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 585 Commercial Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-1024
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,751
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/416,007
; FILING DATE: 04-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06942/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/723-4123
; TELEFAX: 617/723-8962
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-475-751-4

Query Match 15.4%; Score 87; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKFTTSC 114
| | | | | | | | | | | | | | | |
Db 1 AGCKNFFWKFTTSC 14

-----RESULT 14
US-08-282-980B-1
; Sequence 1, Application US/08282980B
; Patent No. 5932189
; GENERAL INFORMATION:
; APPLICANT: Dean, Richard T.
; APPLICANT: McBride, William
; APPLICANT: Lister-James, John
; TITLE OF INVENTION: Peptides
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,980B
; FILING DATE: 29-JUL-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5932189nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 92,385-I
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3..14

```

OTHER INFORMATION: /label- Disulfide bond
OTHER INFORMATION: /note- "A disulfide bond exists between the
OTHER INFORMATION: two sulfur atoms of the cysteine residues;
US-08-282-980B-1

Query Match 15.4%; Score 87; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKFTTSC 114
| | | | | | | | | | | | | | | |
Db 1 AGCKNFFWKFTTSC 14

RESULT 15

US-08-747-137-13
; Sequence 13; Application US/08747137
; Patent No. 5945033

GENERAL INFORMATION:

APPLICANT: YEN, Richard C.K.
TITLE OF INVENTION: NON-CROSSLINKED PROTEIN PARTICLES FOR
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC USE
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/747,137
FILING DATE: 12-NOV-1996
CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/212,546
FILING DATE: 14-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/069,831
FILING DATE: 01-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/959,560
FILING DATE: 13-OCT-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/641,720
FILING DATE: 15-JAN-1991

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 016197-000840US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200

SEQUENCE CHARACTERISTICS:

INFORMATION FOR SEQ ID NO: 13:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
US-08-747-137-13

Query Match 15.4%; Score 87; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKFTTSC 114
| | | | | | | | | | | | | | | |

Db 1 AGCKNFFWKFTTSC 14

Search completed: June 13, 2002, 12:22:55
Job time: 193 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:20:22 ; Search time 35.79 Seconds
(without alignments)
306.068 Million cell updates/sec

Title: US-09-727-739B-3
Perfect score: 566
Sequence: 1 MLSTRVQCALALLSLALAI.....APRERKAGCKNFFWKFTTSC 114

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

otal number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	409	72.3	114	1 RIIDS1	somatostatin-14 pr
2	409	72.3	114	2 I50798	preprosomatostatin
3	400	70.7	116	1 S20630	somatostatin precu
4	382.5	67.6	115	2 JC6166	somatostatin-14 pr
5	360	63.6	116	1 RIHUS1	somatostatin I pre
6	360	63.6	116	1 A28968	somatostatin I pre
7	352	62.2	116	1 RIBOS1	somatostatin precu
8	347	61.3	116	1 RIRTS1	somatostatin precu
9	347	61.3	116	1 RIMSS1	somatostatin precu
10	293	51.8	92	1 RIPGS	somatostatin I pre
11	251	44.3	121	1 RIAFS1	somatostatin I pre
12	177.5	31.4	115	2 I51064	somatostatin II pr
13	162.5	28.7	125	1 RIAFS2	somatostatin II pr
14	127	22.4	28	2 A61322	somatostatin-28 -
15	116.5	20.6	34	2 A32271	somatostatin-relat
16	114	20.1	73	2 S00169	somatostatin II pr
17	110.5	19.5	105	1 RIIDS2	somatostatin-22 pr
18	107.5	19.0	103	2 JC6167	somatostatin-14 [P
19	107	18.9	25	2 A60842	somatostatin-25 -
20	106	18.7	37	2 A32000	somatostatin, panc
21	103.5	18.3	74	2 S00166	somatostatin II pr
22	94.5	16.7	112	2 S67489	cortistatin precu
23	87	15.4	14	2 C60414	somatostatin - sli
24	87	15.4	14	2 B60842	somatostatin I - E
25	87	15.4	14	2 A60840	somatostatin I - E
26	87	15.4	14	2 S00172	somatostatin I - s
27	87	15.4	25	2 B60840	somatostatin-25 -
28	86.5	15.3	105	2 JC5414	cortistatin-like p
29	82	14.5	14	2 A60622	somatostatin - spo

30	79.5	14.0	846	2 S52418	GTP-binding regula
31	77.5	13.7	571	2 AE3281	hypothetical cytos
32	77	13.6	385	2 T18821	hypothetical prote
33	74	13.1	297	2 T48855	probable heat shoc
34	74	13.1	965	2 AE0418	valine--trNA ligas
35	74	13.1	1385	2 T13887	tlr protein - frui
36	74	13.1	1389	2 T13852	gene wheeler prote
37	74	13.1	1732	2 T14039	protein kinase (EC
38	73	12.9	493	2 S49175	legumain (EC 3.4.2
39	72.5	12.8	216	2 T40268	ubiquinone biosynt
40	72.5	12.8	274	2 S72643	type 4 fimbrial bi
41	71.5	12.6	1285	2 H85041	hypothetical prote
42	71	12.5	1295	2 T24587	hypothetical prote
43	70.5	12.5	478	2 AB3641	fumarate hydratase
44	70.5	12.5	925	2 E83529	sensor/response re
45	70	12.4	414	2 T15947	hypothetical prote

ALIGNMENTS

RESULT 1

RIIDS1
somatostatin-14 precursor - channel catfish
N;Alternate names: somatostatin I
N;Contains: somatostatin-14
C;Species: Ictalurus punctatus (channel catfish)
C;Date: 30-Jun-1980 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999
C;Accession: S00292; A93897; A92334; A01435
R;Minth, C.D.; Taylor, W.L.; Magazin, M.; Tavianini, M.A.; Collier, K.; Weith, H.L.;
J. Biol. Chem. 257, 10372-10377, 1982
A;Title: The structure of cloned DNA complementary to catfish pancreatic somatostat.
A;Reference number: S00292; MUID:82265698
A;Accession: S00292
A;Molecule type: mRNA
A;Residues: 1-114 <MIN>
A;Cross-references: EMBL:V00607; NID:g64017; PIDN:CAA3877.1; PID:g64018
R;Taylor, W.L.; Collier, K.J.; Deschenes, R.J.; Weith, H.L.; Dixon, J.E.
Proc. Natl. Acad. Sci. U.S.A. 78, 6694-6698, 1981
A;Title: Sequence analysis of a cDNA coding for a pancreatic precursor to somatostat.
A;Reference number: A93897; MUID:82082515
A;Accession: A93897
A;Molecule type: mRNA
A;Residues: 82-108 <TAY>
A;Cross-references: GB:J00944
R;Andrews, P.C.; Dixon, J.E.
J. Biol. Chem. 256, 8267-8270, 1981
A;Title: Isolation and structure of a peptide hormone predicted from a mRNA sequence
A;Reference number: A92334; MUID:81264223
A;Accession: A92334
A;Molecule type: protein
A;Residues: 101-114 <AND>
C;Superfamily: somatostatin
C;Keywords: neuropeptide
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-100/Domain: propeptide #status predicted <PRO>
F:101-114/Product: somatostatin-14 #status experimental <MAT>
F:103-114/Disulfide bonds: #status experimental

Query Match 72.3%; Score 409; DB 1; Length 114;
Best Local Similarity 69.3%; Pred. No. 4.5e-32;
Matches 79; Conservative 15; Mismatches 20; Indels 0; Gaps 0;

QY 1 MLSTRVQCALALLSLALAISSVSAAPSDAKLRQLQRLSLMAPAGKQELARNTLVLLSEL 60

Db 1 MFSTRIQCALALLAVALSVCSVSGAPSDAKLRQLQRLSILAPSVKQELTRYTLAEJLAEL 60

QY 61 AHVNEAIELDDMSHGVEQEDVDLELERAPGPVLAPRERKAGCKNFFWKFTTSC 114

Db 61 AQAENEVLDSDEVSRAAESEGARLEMERAGPMPLAPRERKAGCKNFFWKFTTSC 114

RESULT 2

I50798

preprosomatostatin SS-14 - channel catfish

C;Species: Ictalurus punctatus (channel catfish)

C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999

C;Accession: I50798

R;Dixon, J.E.; Andrews, P.C.

Adv. Exp. Med. Biol. 188, 19-29, 1985

A;Title: Somatostatins of the channel catfish.

A;Reference number: I50798; MUID:85303576

A;Accession: I50798

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-114 <DIX>

A;Cross-references: GB:M25903; NID:g213339; PIDN:AAA49339.1; PID:g213340

C;Superfamily: somatostatin

Query Match

Best Local Similarity 72.3%; Score 409; DB 2; Length 114;

Matches 79; Conservative 15; Mismatches 20; Indels 0; Gaps 0;

QY

1 MLSTRVQCALALLSLALAISVSVAAPSDAKLRQLLQRLSILMAPAGKQELARNTLVLLSEL 60

1 MPSTRIQCALALLAVALSVCSVSGAPSDAKLRQLRQFLQSLAPSVKQLTRYTLAELLAE 60

QY

61 AHVENEATLEDDMSHGVEQEDVDLELERAPGPVLAPRERKAGCKNFFWKTFTSC 114

61 AEAEDEVLDSEVSRAESEGARLEMERAAAGPMLAPRERKAGCKNFFWKTFTSC 114

RESULT 3

S20630

somatostatin precursor - chicken

C;Species: Gallus gallus (chicken)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

R;Nata, K.; Kobayashi, T.; Karahashi, K.; Kato, S.; Yamamoto, H.; Okamoto,

submitted to the EMBL Data Library, June 1991

A;Description: Nucleotide sequence determination of chicken somatostatin precursor cDNA.

A;Reference number: S20630

A;Accession: S20630

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-116 <NAT>

A;Cross-references: EMBL:X60191; NID:g62985; PIDN:CAA42747.1; PID:g62986.

C;Superfamily: somatostatin

Query Match

Best Local Similarity 70.7%; Score 400; DB 1; Length 116;

Matches 81; Conservative 14; Mismatches 19; Indels 2; Gaps 1;

QY

1 MLSTRVQCALALLSLALAISVSVAAPSDAKLRQLLQRLSILMAPAGKQELARNTLVLLSEL 60

1 MLSCRLOCALALLSIALAVGTVSAAPSDPRLRQFLQSLAAAGKQELAKYFLAELLSEP 60

QY

61 AHVENEATLEDDMSHGVEQEDVDLELERAPGPVLAPRERKAGCKNFFWKTFTSC 114

61 SQTENEALSEDLRGAEQDEVRLERLSANSNPALAPRERKAGCKNFFWKTFTSC 116

RESULT 4

JC6166

somatostatin-14 precursor - laughing frog

N;Alternate names: PSS1 protein

C;Species: Rana ridibunda (laughing frog)

C;Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 16-Jul-1999

C;Accession: JC6166

R;Tostivint, H.; Lohrmann, I.; Bucharies, C.; Vileau, D.; Coulouarn, Y.; Fournier, A.; Co

Proc. Natl. Acad. Sci. U.S.A. 93, 12605-12610, 1996

A;Title: Occurrence of two somatostatin variants in the frog brain: Characterization of

A;Reference number: JC6166; MUID:97057290

A;Contents: brain

A;Accession: JC6166

A;Molecule type: mRNA

A;Residues: 1-115 <TOS>

A;Cross-references: GB:U68136; NID:g1890650; PIDN:AAC60093.1; PID:g1890651

C;Comment: This protein acts both as a neurotransmitter/neuromodulator and a hormone.

C;Genetics:

A;Gene: pssl

C;Superfamily: somatostatin

C;Keywords: brain; hormone

Query Match

Best Local Similarity 67.6%; Score 382.5; DB 2; Length 115;

Matches 80; Conservative 13; Mismatches 20; Indels 3; Gaps 2;

QY

1 MLSTRVQCALALLSLALAISVSVAAPSDAKLRQLLQRLSILMAPAGKQELARNTLVLLSEL 60

1 MQSCRVCALTLTLLSLALAINSISAAPDPRLRQFLQKSL-ASAGKQELAKYFLAELLSEP 59

QY

61 AHVENEATLEDDMSHGVEQEDVDLELERAPGPVLAPRERKAGCKNFFWKTFTSC 114

60 SQTENEALSEDDLPRGAEQDEVRLERLSANSNPALAPRERKAGCKNFFWKTFTSC 115

RESULT 5

RIHUS1

somatostatin I precursor - human

N;Alternate names: preprosomatostatin

C;Species: Homo sapiens (man)

C;Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 18-Jun-1999.

C;Accession: A43614; A01430; S09381; S50024

R;Shen, L.P.; Rutter, W.J.

Science 224, 168-171, 1984

A;Title: Sequence of the human somatostatin I gene.

A;Reference number: A43614; MUID:84146798

A;Accession: A43614

A;Molecule type: DNA

A;Residues: 1-116 <SH2>

A;Cross-references: GB:J00306; NID:g338287; PIDN:AAA60566.1; PID:g338288

R;Shen, L.P.; Pictet, R.L.; Rutter, W.J.

Proc. Natl. Acad. Sci. U.S.A. 79, 4575-4579, 1982

A;Title: Human somatostatin I: sequence of the cDNA.

A;Reference number: A01430; MUID:83014931

A;Accession: A01430

A;Molecule type: mRNA

A;Residues: 1-116 <SHE>

A;Cross-references: GB:J00306; NID:g338287; PIDN:AAA60566.1; PID:g338288

A;Experimental source: pancreatic somatostatinoma

R;Gomez, S.; Boileau, G.; Zollinger, L.; Nault, C.; Rholam, M.; Cohen, P.

EMBO J. 8, 2911-2916, 1989

A;Title: Site-specific mutagenesis identifies amino acid residues critical in prohorm

A;Reference number: S09381; MUID:90059875

A;Accession: S09381

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 88-108 <GOM>

R;Odum, L.; Johnsen, A.H.

Biochem. J. 303, 263-268, 1994

A;Title: Human seminal plasma contains somatostatin-64.

A;Reference number: S50024; MUID:95031969

A;Accession: S50024

A;Status: preliminary

A;Molecule type: protein

A;Residues: 53-62;67-82 <ODU>

C;Comment: Somatostatin inhibits the release of somatotropin.

C;Genetics:

A;Gene: GDB:SS

A;Cross-references: GDB:119604; OMIM:182450

A;Map position: 3q28-3q28

A;Introns: 46/3

C;Function:

A;Description: inhibits the secretion of a number of peptide hormones, including somatostatin
C;Superfamily: somatostatin
C;Keywords: hormone; hypothalamus; neuropeptide
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-88/Domain: propeptide #status predicted <PRO>
F;89-116/Product: somatostatin-28 #status predicted <M28>
F;103-116/Product: somatostatin-14 #status predicted <M14>
F;105-116/Disulfide bonds: #status experimental

Query Match 63.6%; Score 360; DB 1; Length 116;
Best Local Similarity 62.1%; Pred. No. 2.2e-27;
Matches 72; Conservative 17; Mismatches 25; Indels 2; Gaps 1;

QY 1 MLSTRVQCALALLSLALAISSVSAAPSDAKLRQLQRLQSLMAPAGKQELARNLTVLLSEL 60
||| |:||||| ||: ||: ||| |:|| ||: ||| |:|||||: | |||||
Db 1 MLSCRLOCALAALSIVLALGCVTGAPSDPRLQFLQKSLAAAGKQELAKYFLAELLSEP 60

QY 61 AHVENEAIELDDMSHGVEQEDVDLELERA--PGPVLAPRERKAGCKNFFWKFTTSC 114
||:|:|:|:|:| ||::| ||::| |:||||| ||||| |||||
b 61 NOTENDALEPEDLSQAAEQDEMRLQSLANSNPAMPAPRERKAGCKNFFWKFTTSC 116

RESULT 6
A28968
somatostatin I precursor - crab-eating macaque
N;Alternate names: preprosomatostatin
N;Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 30-Jun-1989 #sequence_revision 31-Jan-1997 #text_change 18-Jun-1999
C;Accession: A28968
R;Travis, G.H.; Sutcliffe, J.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 1696-1700, 1988
A;Title: Phenol emulsion-enhanced DNA-driven subtractive cDNA cloning: isolation of low-
A;Reference number: A28968; MUID:88144503
A;Accession: A28968
A;Molecule type: mRNA
A;Residues: 1-116 <TRA>
A;Cross-references: GB:M19318; NID:9342298; PIDN:AAA36908.1; PID:9342299
C;Comment: Somatostatin inhibits the release of somatotropin.
C;Superfamily: somatostatin
C;Keywords: hormone; neuropeptide
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-88/Domain: propeptide #status predicted <PRO>
F;89-116/Product: somatostatin-28 #status predicted <M28>
F;103-116/Product: somatostatin-14 #status predicted <M14>
F;105-116/Disulfide bonds: #status predicted

Query Match 63.6%; Score 360; DB 1; Length 116;
Best Local Similarity 62.1%; Pred. No. 2.2e-27;
Matches 72; Conservative 17; Mismatches 25; Indels 2; Gaps 1;

QY 1 MLSTRVQCALALLSLALAISSVSAAPSDAKLRQLQRLQSLMAPAGKQELARNLTVLLSEL 60
||| |:||||| ||: ||: ||| |:|| ||: ||| |:|||||: | |||||
Db 1 MLSCRLOCALAALSIVLALGCVTGAPSDPRLQFLQKSLAAAGKQELAKYFLAELLSEP 60

QY 61 AHVENEAIELDDMSHGVEQEDVDLELERA--PGPVLAPRERKAGCKNFFWKFTTSC 114
||:|:|:|:|:| ||::| ||::| |:||||| ||||| |||||
Db 61 NOTENDALEPEDLSQAAEQDEMRLQSLANSNPAMPAPRERKAGCKNFFWKFTTSC 116

RESULT 7
RIBOS1
somatostatin precursor - bovine
N;Alternate names: preprosomatostatin
N;Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)
C;Species: Bos primigenius taurus (cattle)
C;Date: 06-Mar-1992 #sequence_revision 31-Jan-1997 #text_change 18-Jun-1999
C;Accession: A40929
R;Su, C.J.; White, J.W.; Li, W.H.; Luo, C.C.; Frazier, M.L.; Saunders, G.F.; Chan, L.
Mol. Endocrinol. 2, 209-216, 1988
A;Title: Structure and evolution of somatostatin genes.

A;Reference number: A40929; MUID:88288237
A;Accession: A40929
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-116 <SDA>
A;Cross-references: GB:M31217; NID:9163636; PIDN:AAA30744.1; PID:9163637
A;Note: the authors translated the codon ATT for residue 65 as Asn
C;Comment: Somatostatin inhibits the release of somatotropin.
C;Superfamily: somatostatin
C;Keywords: hormone; neuropeptide
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-88/Domain: propeptide #status predicted <PRO>
F;89-116/Product: somatostatin-28 #status predicted <M28>
F;103-116/Product: somatostatin-14 #status predicted <M14>
F;105-116/Disulfide bonds: #status predicted

Query Match 62.2%; Score 352; DB 1; Length 116;
Best Local Similarity 61.2%; Pred. No. 1.3e-26;
Matches 71; Conservative 17; Mismatches 26; Indels 2; Gaps 1;

QY 1 MLSTRVQCALALLSLALAISSVSAAPSDAKLRQLQRLQSLMAPAGKQELARNLTVLLSEL 60
||| |:||||| ||: ||: ||| |:|| ||: ||| |:|||||: | |||||
Db 1 MLSCRLOCALAALSIVLALGCVTGAPSDPRLQFLQKSLAAAGKQELAKYFLAELLSEP 60

QY 61 AHVENEAIELDDMSHGVEQEDVDLELERA--PGPVLAPRERKAGCKNFFWKFTTSC 114
||:|:|:|:|:| ||::| ||::| |:||||| ||||| |||||
Db 61 NOTENDALEPEDLSQAAEQDEMRLQSLANSNPAMPAPRERKAGCKNFFWKFTTSC 116

RESULT 8
RIRTS1

somatostatin precursor - rat
N;Alternate names: preprosomatostatin
N;Contains: somatostatin-14; somatostatin-28
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 18-Jun-1999
C;Accession: A20983; A01431; A47598; A22529; I55220; I51829
R;Montminy, M.R.; Goodman, R.H.; Horovitch, S.J.; Habener, J.F.
Proc. Natl. Acad. Sci. U.S.A. 81, 3337-3340, 1984
A;Title: Primary structure of the gene encoding rat preprosomatostatin.
A;Reference number: A20983; MUID:84221954
A;Accession: A20983
A;Molecule type: DNA
A;Residues: 1-116 <MON>
A;Cross-references: GB:J00787; NID:9207024; PIDN:AAA42164.1; PID:9207025
A;Note: the authors translated the codon ACC for residue 43 as Tyr
R;Argos, P.; Taylor, W.L.; Minth, C.D.; Dixon, J.E.
J. Biol. Chem. 258, 8788-8793, 1983
A;Title: Nucleotide and amino acid sequence comparisons of preprosomatostatins.
A;Reference number: A01431; MUID:83238516
A;Accession: A01431
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-116 <ARG>
R;Benolt, R.; Ling, N.; Esch, F.
Science 238, 1126-1129, 1987
A;Title: A new prosomatostatin-derived peptide reveals a pattern for prohormone clea
A;Reference number: A47598; MUID:88070564
A;Accession: A47598
A;Status: preliminary
A;Molecule type: protein
A;Residues: 25-34 <BEN>
R;Tavianini, M.A.; Hayes, T.E.; Magazin, M.D.; Minth, C.D.; Dixon, J.E.
J. Biol. Chem. 259, 11798-11803, 1984
A;Title: Isolation, characterization, and DNA sequence of the rat somatostatin gene.
A;Reference number: A22529; MUID:85006903
A;Accession: A22529
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-116 <TAV>
A;Cross-references: GB:K02248; NID:9207014; PIDN:AAA42161.1; PID:9207017
R;Goodman, R.H.; Jacobs, J.W.; Dee, P.C.; Habener, J.F.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:22:28 ; Search time 18.73 Seconds
(without alignments)
235.666 Million cell updates/sec

Title: US-09-727-739B-3
Perfect score: 566
Sequence: 1 MLSTRVQCALALLSLALIS.....APRERKAGCKNFFWKFTTSC 114

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	ID	Description
1	456	80.6	114	1 SMSA_CARAU	Q9ygh5 carassius a
2	409	72.3	114	1 SMS1_ICTPU	P01171 ictalurus p
3	400	70.7	116	1 SMS_CHICK	P33094 gallus gall
4	382.5	67.6	115	1 SMS1_RANRI	P87384 rana ridibu
5	373.5	66.0	115	1 SMS1_PROAN	Q9w7f0 prototenus
6	361	63.8	116	1 SMS_CANFA	P49670 canis fami
7	361	63.8	116	1 SMS_SHEEP	O46688 ovis aries
8	360	63.6	116	1 SMS_HUMAN	P01166 homo sapien
9	352	62.2	116	1 SMS_BOVIN	P26917 bos taurus
10	347	61.3	116	1 SMS_MOUSE	P01167 mus musculu
11	293	51.8	92	1 SMS_PIG	P01168 sus scrofa
12	251	44.3	121	1 SMS1_LOPAM	P01169 lophius ame
13	177.5	31.4	115	1 SMS2_ONCMY	Q91194 oncorhynchu
14	163.5	28.9	125	1 SMS2_LOPAM	P01170 lophius ame
15	153	27.0	120	1 SMS2_CARAU	Q9ygh4 carassius a
16	129	22.8	26	1 SMS1_AMICA	Q9prz6 amia calva
17	123	21.7	111	1 SMSB_CARAU	Q9ygh3 carassius a
18	117	20.7	109	1 SMS2_PROAN	Q9w7e9 prototenus
19	116.5	20.6	34	1 SMS_MYXGL	P19209 myxine glut
20	114	20.1	73	1 SMS2_PLAFE	P21780 platichthys
21	110.5	19.5	105	1 SMS2_ICTPU	P01172 ictalurus p
22	108	19.1	35	1 SMS_LAMFL	Q9prro lampetra fl
23	107.5	19.0	103	1 SMS2_RANRI	P87385 rana ridibu
24	106	18.7	37	1 SMS_PETMA	P21779 petromyzon
25	103.5	18.3	74	1 SMS2_MYOSC	P09876 myoxocephal
26	99	17.5	28	1 SMS2_ORENI	P81029 oreochromis
27	94.5	16.7	112	1 CORT_RAT	Q62949 rattus norv
28	87	15.4	14	1 SMS1_MYOSC	P20750 myoxocephal
29	87	15.4	14	1 SMS_ALLMI	P31885 alligator m
30	86.5	15.3	105	1 CORT_HUMAN	O00230 homo sapien
31	85	15.0	109	1 CORT_MOUSE	P56469 mus musculu
32	74	13.1	297	1 HTPX_STRGC	O30795 streptococc
33	73	12.9	493	1 VPE_VICSA	P49044 vicia sativ

34	72	12.7	256	1	CPI_ELAUM	O65333 elaeagnus u
35	72	12.7	1436	1	MRP5_RAT	Q9qym0 rattus norv
36	70.5	12.5	1087	1	AKA9_RABIT	Q28628 oryctolagus
37	70.5	12.5	3067	1	CALC_MOUSE	Q60847 mus musculu
38	69.5	12.3	616	1	MUTA_STRCM	Q05064 streptomyc
39	69	12.2	98	1	SNC5_HUMAN	O75971 homo sapien
40	69	12.2	402	1	OPS4_CANAL	P46596 candida alb
41	69	12.2	453	1	SSF2_YEAST	Q12153 saccharomyc
42	69	12.2	646	1	NODQ_RHIS3	P72339 r nodq bifu
43	68.5	12.1	442	1	MFAL_CHICK	P55080 gallus gall
44	68	12.0	327	1	FRA2_RAT	P51145 rattus norv
45	68	12.0	396	1	Y858_TREPA	O83830 treponema p

ALIGNMENTS

RESULT 1

SMSA_CARAU

ID SMSA_CARAU STANDARD; PRT; 114 AA.

AC Q9YGH5;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Somatostatin IA precursor [Contains: Somatostatin-26; Somatostatin-14]

OS Carassius auratus (Goldfish)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;

OC Cypriniformes; Cyprinidae; Carassius

OX NCBI_TaxID=7957;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Lin X.-W., Peter R.E.;

RT "Cloning and characterization of cDNAs encoding preprosomatostatin-I

RT and -II from goldfish brain."

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

CC -I- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.

CC -I- SUBCELLULAR LOCATION: Secreted.

CC -I- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U40754; AAD09359.1;

DR InterPro; IPR004250; Somatostatin.

DR Pfam; PF03002; Somatostatin; 1.

KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.

FT SIGNAL 1 24 POTENTIAL.

FT PROPEP 25 88 POTENTIAL.

FT PEPTIDE 89 114 SOMATOSTATIN-26 (POTENTIAL).

FT PEPTIDE 101 114 SOMATOSTATIN-14.

FT DISULFID 103 114 BY SIMILARITY.

SQ SEQUENCE 114 AA; 12574 MW; B5920015E2D272A4 CRC64;

Query Match 80.6%; Score 456; DB 1; Length 114;
Best Local Similarity 78.9%; Pred. No. 2.1e-36;
Matches 90; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY 1 MLSTRVQCALALLSLALISVSAAPSADKLRLQLRSLMAPAGKQLARNTLVLLSEL 60

DB 1 MLSTRIQCALALLSLALAVCSVAAPTDAKLRLQLRSLNPNAGKQLARYTLADLLSEL 60

QY 61 AHVENEAIELDDMSHGVEQEDVDLELERAPGVPLAPRERKAGCKNFFWKFTTSC 114

DB 61 VQAEAELEPEDLSRAVEKDEVRLERAEAGMPLAPRERKAGCKNFFWKFTTSC 114

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Somatostatin I precursor (PSS1) [Contains: Somatostatin-14 (S-I) (SSS1)]
 OS Rana ridibunda (Laughing frog) (Marsh frog)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
 OX NCBI_TaxID=8406;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97057290; PubMed=8901629;
 RA Tostivint H., Lihmann I., Buchard C., Vieau D., Coulouarn Y., Fournier A., Conlon J.M., Vaudry H.;
 RT "Occurrence of two somatostatin variants in the frog brain: characterization of the cDNAs, distribution of the mRNAs, and receptor-binding affinities of the peptides.";
 RT Proc. Natl. Acad. Sci. U.S.A. 93:12605-12610(1996).
 RL [2]
 Y SEQUENCE OF 102-115.
 RC TISSUE=Brain;
 RX MEDLINE=93038702; PubMed=1358069;
 RA Vaudry H., Chartrel N., Conlon J.M.;
 RT "Isolation of [Pro2, Met13]somatostatin-14 and somatostatin-14 from the frog brain reveals the existence of a somatostatin gene family in a tetrapod.";
 RT Biochem. Biophys. Res. Commun. 188:477-482(1992).
 RL Biochem. Biophys. Res. Commun. 188:477-482(1992).
 CC -|- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U68136; AAC60093.1;
 DR HSSP; P03619; 3PAL.
 DR InterPro; IPR004250; Somatostatin.
 DR Pfam; PF03002; Somatostatin; 1.
 KW Cleavage on pair of basic residues; Hormone; Multigene family; Signal.
 FT SIGNAL 1 24 BY SIMILARITY.
 FT PROPEP 25 99 BY SIMILARITY.
 FT PEPTIDE 102 115 SOMATOSTATIN-14.
 FT DISULFID 104 115 BY SIMILARITY.
 SQ SEQUENCE 115 AA; 12691 MW; 349756FEB4ABE213 CRC64;

Query Match 67.6%; Score 382.5; DB 1; Length 115;
 Best Local Similarity 69.0%; Pred. No. 1.8e-29;
 Matches 80; Conservative 13; Mismatches 20; Indels 3; Gaps 2;
 QY 1 MLSTRVQCALALLSLAALSSVSAAPSDAKLRQLQSLMAPAGKQELARNTLVLLSEL 60
 Db 1 MQSCRVQCALTLALLSLAALSSVSAAPSDAKLRQLQSLMAPAGKQELARNTLVLLSEL 59
 QY 61 AHVNEAIEIDDMSHGVEQEDVDLELER--PGPVLAPRRERKAGCKNFFWKTFTSC 114
 Db 60 SQTDAEALSDLPFGAEQDEVRLELERSANSPPALAPRRERKAGCKNFFWKTFTSC 115
 RESULT 5
 SMSL_PROAN STANDARD; PRT; 115 AA.
 ID SMSL_PROAN
 AC Q9W7F0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Somatostatin I precursor (PSS1) [Contains: Somatostatin-27;

Somatostatin-14].
 OS Protopterus annectens (African lungfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Dipnoi; Lepidosireniformes; Protopterygidae; Protopterus.
 OX NCBI_TaxID=7888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99326690; PubMed=10398054;
 RA Trabucchi M., Tostivint H., Lihmann I., Jegou S., Vallarino M., Vaudry H.;
 RT "Molecular cloning of the cDNAs and distribution of the mRNAs encoding two somatostatin precursors in the African lungfish Protopterus annectens.";
 RT J. Comp. Neurol. 410:643-652(1999).
 CC -|- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF126243; AAD39138.1;
 DR InterPro; IPR004250; Somatostatin.
 DR Pfam; PF03002; Somatostatin; 1.
 KW Cleavage on pair of basic residues; Hormone; Multigene family; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT PROPEP 25 88 POTENTIAL.
 FT PEPTIDE 89 115 SOMATOSTATIN-27 (POTENTIAL).
 FT PEPTIDE 102 115 SOMATOSTATIN-14.
 FT DISULFID 104 115 BY SIMILARITY.
 SQ SEQUENCE 115 AA; 12600 MW; BOCEFE1603FEAF09 CRC64;
 Query Match 66.0%; Score 373.5; DB 1; Length 115;
 Best Local Similarity 67.8%; Pred. No. 1.2e-28;
 Matches 78; Conservative 12; Mismatches 24; Indels 1; Gaps 1;
 QY 1 MLSTRVQCALALLSLAALSSVSAAPSDAKLRQLQSLMAPAGKQELARNTLVLLSEL 60
 Db 1 MLSCRFQCALVLLSLAVFESKVSAPSDAKLRQLQSLMAPAGKQELARNTLVLLSEL 60
 QY 61 AHVNEAIEIDDMSHGVEQEDVDLELER--PGPVLAPRRERKAGCKNFFWKTFTSC 114
 Db 61 AQSENDALSDLSRGADQDEVRLELERSANSPPALAPRRERKAGCKNFFWKTFTSC 115
 RESULT 6
 SMS_CANFA STANDARD; PRT; 116 AA.
 ID SMS_CANFA
 AC P49670;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].
 GN SST.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Gastric mucosa;
 RX MEDLINE=97142297; PubMed=8988514;
 RA Dickinson C.J., Delvalle J., Todisco A., Gantz I., Tong L., Finniss S., Yamada T.;
 RT "Canine prosomatostatin: isolation of a cDNA, regulation of gene expression, and characterization of post-translational processing


```
CC EMBL; U32471; AAC59695.1;
DR InterPro: IPR004250; Somatostatin.
DR Pfam: PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
FT SIGNAL 1 18 POTENTIAL.
FT PROPEP 19 87 POTENTIAL.
FT PEPTIDE 88 115 [TYR21, GLY24] SOMATOSTATIN-28 (POTENTIAL).
FT PEPTIDE 102 115 [TYR7, GLY10] SOMATOSTATIN-14.
FT DISULFID 104 115 BY SIMILARITY.
SQ SEQUENCE 115 AA; 12963 MW; 520595025FCA6D91 CRC64;

Query Match 31.4%; Score 177.5; DB 1; Length 115;
Best Local Similarity 34.7%; Pred. No. 3.6e-10;
Matches 50; Conservative 9; Mismatches 18; Indels 67; Gaps 4;

QY 5 RVQCALALLSLALAISSVSAAPS---DAKLRLQLRSLMA----- 41
Db 5 RIHCALALLGLALAISSQGAASQPDLDLRSRLRLQRAAAWPHRSGVSEWRWTFYPCNP 64
Y 42 -----PAGKQELARNTLVLELLSELAHVENEAIELDDMSHGVEQEDVDLELERAPGP 92
Db 65 CLRPRKVKCPAG-----AKEDLRVELERSVGN 91

QY 93 V--LAPRERKAGCKNFFWKTFTSC 114
Db 92 PNNLPPEKAGCKNFFWKTFTSC 115

RESULT 14
SMS2_LOPAM STANDARD; PRT; 125 AA.
AC P01170; Q91066;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr7, Gly10]somatostatin-14].
OS Lophius americanus (American goosefish) (Anglerfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Lophiiformes; Lophiidae; Lophius.
OX NCBI_TaxID=8073;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81052423; PubMed=6107860;
RA Hobart P.M., Crawford R., Shen L., Pictet R., Rutter W.J.;
RT "Cloning and sequence analysis of cDNAs encoding two distinct
somatostatin precursors found in the endocrine pancreas of
anglerfish.";
RL Nature 288:137-141(1980).
RN [2]
RP PARTIAL SEQUENCE, AND HYDROXYLATION.
RX MEDLINE=87308304; PubMed=2887572;
RA Andrews P.C., Nichols R., Dixon J.E.;
RT "Post-translational processing of preprosomatostatin-II examined
using fast atom bombardment mass spectrometry.";
RL J. Biol. Chem. 262:12692-12699(1987).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ACTIVITY OR A DIFFERENT TYPE OF TARGET CELL FROM SOMATOSTATIN I.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V00641; CAA23987.1; -
```

```
DR PIR; A01434; RIAFS2.
DR InterPro: IPR004250; Somatostatin.
DR Pfam: PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal; Hydroxylation;
KW Multigene family.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 109 [TYR7, GLY10] SOMATOSTATIN-14.
FT PEPTIDE 112 125 [TYR7, GLY10] SOMATOSTATIN-14.
FT DISULFID 114 125 HYDROXYLATION.
FT MOD_RES 120 120
FT CONFLICT 77 78 DV -> TG (IN REF. 1).
FT CONFLICT 90 90 G -> E (IN REF. 1).
SQ SEQUENCE 125 AA; 14052 MW; 5E14605D7B9A46FE CRC64;

Query Match 28.9%; Score 163.5; DB 1; Length 125;
Best Local Similarity 38.2%; Pred. No. 8.2e-09;
Matches 50; Conservative 17; Mismatches 41; Indels 23; Gaps 5;

QY 1 MLSTRVQCALALLSLALAISSVSA-----APSDAKLRQ--LLQRLMAPAGKQELA 49
Db 1 MQCIRCPAILALLALVLCGSPVSSQLDREQSDNQDLDLELRQHLLERARSAGLLSQEWS 60
QY 50 RNTLVLELLSEL---AHVENEAIELDDMSHGVEQEDVDLELERAPGPV--LAPRERKAGC 103
Db 61 KRAVEELLAQMSLPEDVQREAEADASMATGG-----RMNLERSVDSTNNLPPRERKAGC 114
QY 104 KNFFWKFTFTSC 114
Db 115 KNFYWKGTFTSC 125

RESULT 15
SMS2_CARAU STANDARD; PRT; 120 AA.
AC Q9YGH4; Q9PTU2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr21, Gly24]somatostatin-28;
[Tyr7, Gly10]somatostatin-14].
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Lin X.-W., Peter R.E.;
RT "Cloning and characterization of cDNAs encoding preprosomatostatin-I
and -II from goldfish brain.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Otto C.J., Peter R.E.;
RT "The expression of SRIF mRNA in the brain of goldfish.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U60262; AAD09626.1; -
```

DR	EMBL; AF025686; AAF15306.1; -
DR	InterPro; IPR004250; Somatostatin.
DR	Pfam; PF03002; Somatostatin; 1.
KW	Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
FT	SIGNAL 1 23 POTENTIAL.
FT	PROPEP 24 92 POTENTIAL.
FT	PEPTIDE 93 120 [TYR21, GLY24]SOMATOSTATIN-28.
FT	PEPTIDE 107 120 [TYR7, GLY10]SOMATOSTATIN-14.
FT	DISULFID 109 120 BY SIMILARITY.
FT	CONFLICT 51 51 Q -> RW (IN REF. 2).
SQ	SEQUENCE 120 AA; 13723 MW; 98957D68011A651A CRC64;

Query Match	27.0%;	Score 153;	DB 1;	Length 120;
Best Local Similarity	37.9%;	Pred. No. 7.6e-08;		
Matches 44;	Conservative 16;	Mismatches 48;	Indels 8;	Gaps 3;

[illegible]

Search completed: June 13, 2002, 12:29:53
Job time: 445 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:22:58 ; Search time 60.72 Seconds
(without alignments)
324.793 Million cell updates/sec

Title: .
 US-09-727-739B-3
 perfect score: 566
 Sequence: 1 MLSTRVQCALLES LALAI. APRERKAGCKNFFWKTFSC 114

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :
SPTRMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archheap:*
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	456	80.6	114	13	Q90Y41	Q90Y41 gnathonemus
2	454	80.2	114	13	Q90Y40	Q90Y40 chitala chi
3	400	70.7	116	13	Q90XE1	Q90XE1 acipenser t
4	208	36.7	114	13	Q90Y42	Q90Y42 pantodon bu
5	201.5	35.6	115	13	Q90Y43	Q90Y43 osteoglossu
6	196	34.6	120	13	Q90Y39	Q90Y39 catostomus
7	117	20.7	111	13	Q90XE0	Q90XE0 acipenser t
8	107.5	19.0	107	13	Q9DDE4	Q9DDE4 brachydanio
9	99	17.5	28	13	Q9PRN9	Q9PRN9 carassius a
10	87	15.4	25	13	Q9PRV0	Q9PRV0 anguilla ja
11	85	15.0	105	11	Q9RLP8	Q9RLP8 mus musculus
12	79.5	14.0	715	11	Q63803	Q63803 rattus norv
13	77	13.6	376	5	O01392	O01392 caenorhabdi
14	77	13.6	385	5	Q17355	Q17355 caenorhabdi
15	75	13.3	323	5	Q9BJ45	Q9BJ45 leishmania
16	74.5	13.2	422	11	Q9JJX0	Q9JJX0 mus musculus

17	74.5	13.2	756	11	Q921N8	Q9z1n8 mus musculus
18	74	13.1	557	5	Q961H0	Q961h0 drosophila
19	74	13.1	1385	5	Q9V8Z5	Q9v8z5 drosophila
20	74	13.1	1389	5	Q24591	Q24591 drosophila
21	74	13.1	1732	11	O54874	O54874 rattus norv
22	73	12.9	499	10	Q94CT1	Q94ct1 oryza sativ
23	73	12.9	1021	4	Q96M86	Q96m86 homo sapien
24	72.5	12.8	216	3	O74826	O74826 schizosacch
25	72.5	12.8	274	16	O51534	O51534 pseudomonas
26	72.5	12.8	804	5	Q9NJA3	Q9nja3 aedes aegyp
27	72	12.7	299	2	Q93D93	Q93d93 streptococc
28	71.5	12.6	284	2	Q9KZP7	Q9kzp7 streptomyce
29	71.5	12.6	1160	5	Q9VJH2	Q9vjh2 drosophila
30	71.5	12.6	1285	10	Q9ZR01	Q9zr01 arabidopsis
31	71	12.5	296	16	Q98CQ4	Q98cq4 rhizobium l
32	71	12.5	386	17	Q973P4	Q973p4 sulfolobus
33	71	12.5	412	5	Q9V8D0	Q9v8d0 drosophila
34	71	12.5	995	10	O22974	O22974 arabidopsis
35	71	12.5	1295	5	Q22257	Q22257 caenorhabdi
36	70.5	12.5	814	2	O52767	O52767 shewanella
37	70.5	12.5	925	16	O31138	O31138 pseudomonas
38	70	12.4	414	5	Q19088	Q19088 caenorhabdi
39	70	12.4	437	16	Q927Y9	Q927y9 listeria in
40	70	12.4	480	2	Q9LAX3	Q9lax3 streptococc
41	70	12.4	645	16	Q985Q4	Q985q4 rhizobium l
42	70	12.4	1030	10	Q9AWS4	Q9aws4 oryza sativ
43	70	12.4	1038	10	Q94CN3	Q94cn3 oryza sativ
44	69.5	12.3	332	16	Q9HZ43	Q9hz43 pseudomonas
45	69.5	12.3	386	10	O49617	O49617 arabidopsis

ALIGNMENTS

RESULT 1
Q90Y41
ID Q90Y41 PRELIMINARY; PRT; 114 AA.
AC Q90Y41;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

Query Match	80.6%	Score 456;	DB 13;	Length 114;
Best Local Similarity	78.9%;	Pred. NO. 8.1e-37;		
Matches 90;	Conservative 11;	Mismatches 13;	Indels 0;	Gaps 0;

[illegible]

QY : 61 AHVNEAIELDDMSHGVEQEEDVDLELERAGPVLAPRRKAGCKNFFWKFTTSC 114
 |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
 Db: 61 VRVNEALEPPDLSRGADOOEVRILEERAAGPATAPRRKAGCKNFYWKGTFTSC 114

RESULT 2
Q90Y40

```

ID Q90Y40 PRELIMINARY; PRT; 114 AA.
AC Q90Y40;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Chitalla chitalla.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Notopteridae; Chitalla.
OX NCBI_TaxID=112163;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292653; AAK97070.1; -
SQ SEQUENCE 114 AA; 12561 MW; 4E3C32F58E34F971 CRC64;

Query Match 80.2%; Score 454; DB 13; Length 114;
Best Local Similarity 79.8%; Pred. No. 1.3e-36;
Matches 91; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 MLSTRVOCALALLSLALAISSVSAAPSDAKLRQLLQSLMAPAGKQELARNTLVLLSEL 60
Db 1 MLSTRVOCALALLSLALAISSVSAAPSDAKLRQLLQSLMAPAGKQELARNTLVLLSEL 60

QY 61 AHVENEALDDMSHGVQEDVDLELERAPGPVLAPRERKAGCKNFFWKTFTSC 114
Db 61 AHVENEALDDMSHGVQEDVDLELERAPGPVLAPRERKAGCKNFFWKTFTSC 114

RESULT 3
Q90XEL PRELIMINARY; PRT; 116 AA.
ID Q90XEL;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SOMATOSTATIN.
OS Acipenser transmontanus (White sturgeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenserinae; Acipenser.
OX NCBI_TaxID=7904;
RN [1]
RP SEQUENCE FROM N.A.
RA Trabucchi M., Tostivint H., Lihrmann I., Sollars C., Vallarino M.,
RA Dores R.M., Vaudry H.;
RT "Polygenic expression of somatostatin in the sturgeon Acipenser
RT transmontanus: molecular cloning and distribution of the mRNAs
RT encoding two somatostatin precursors.";
RL J. Comp. Neurol. 0:0-0(2001).
DR EMBL; AF395849; AAL13248.1; -
SQ SEQUENCE 116 AA; 12616 MW; 72E0C3FF6C80650F CRC64;

```

```

Query Match 70.7%; Score 400; DB 13; Length 116;
Best Local Similarity 69.8%; Pred. No. 2.2e-31;
Matches 81; Conservative 14; Mismatches 19; Indels 2; Gaps 1;

QY 1 MLSTRVOCALALLSLALAISSVSAAPSDAKLRQLLQSLMAPAGKQELARNTLVLLSEL 60
Db 1 MLSTRVOCALALLSLALAISSVSAAPSDAKLRQLLQSLMAPAGKQELARNTLVLLSEL 60

QY 61 AHVENEALDDMSHGVQEDVDLELERAP--GPVLAPRERKAGCKNFFWKTFTSC 114
Db 61 AHVENEALDDMSHGVQEDVDLELERAP--GPVLAPRERKAGCKNFFWKTFTSC 114

```

```

RESULT 4
Q90Y42 PRELIMINARY; PRT; 114 AA.
ID Q90Y42;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Pantodon buchholzi (Butterflyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Pantodontidae; Pantodon.
OX NCBI_TaxID=8276;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292651; AAK97068.1; -
SQ SEQUENCE 114 AA; 12352 MW; 7E3D44CB6A27B12F CRC64;

```

```

Query Match 36.7%; Score 208; DB 13; Length 114;
Best Local Similarity 48.7%; Pred. No. 8.5e-13;
Matches 55; Conservative 14; Mismatches 38; Indels 6; Gaps 5;

QY 5 RVOCALALLSLALAISSVSAAPSDAKLRQLLQSLMAPAGKQELARNTLVLLSEL 63
Db 5 QVHCILALLGLVLGMCSSSATQLDSRYSLVQRRARASMPQDWGKLS-VEELSLLAAT 63

QY 64 ENEAIELDDMSHGVQEDVDLELERAPGPVLAPRERKAGCKNFFWKTFTSC 114
Db 64 EAD-MPEGDMSAEESGAHLDLERSVEFGNV-PPRERKAGCKNFFWKTFTSC 114

```

```

RESULT 5
Q90Y43 PRELIMINARY; PRT; 115 AA.
ID Q90Y43;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Osteoglossum bicirrhosum (silver arowana).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
OX NCBI_TaxID=109271;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292650; AAK97067.1; -
SQ SEQUENCE 115 AA; 12791 MW; D65FBD7C6F1E4E4D CRC64;

```

```

Query Match 35.6%; Score 201.5; DB 13; Length 115;
Best Local Similarity 43.8%; Pred. No. 3.7e-12;
Matches 49; Conservative 17; Mismatches 43; Indels 3; Gaps 3;

QY 5 RVOCALALLSLALAISSVSAAPSDAKLRQLLQSLMAPAGKQELARNTLVLLSEL 63
Db 5 QIHCTLVLLGLVLGILYCPASQPDRLRYRSLVQRRARASMPQDWGKLS-VEELSLLAAT 64

QY 64 ENEAIELDDMSHGVQEDVDLELERAPGPVLAPRERKAGCKNFFWKTFTSC 114
Db 65 QGE-VPOGAVSADEEDVRVLDLERSLELNLPPRERKAGCKNFFWKTFTSC 115

```

```

RESULT      6
Q90Y39      PRELIMINARY;          PRT;    120 AA.
ID   Q90Y39
AC   Q90Y39;
DT   01-DEC-2001 (TREMBLrel. 19, Created)
DT   01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE   PREPROSOMATOSTATIN.
OS   Catostomus commersoni (White sucker).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Ostariophysi;
OC   Cypriniformes; Catostomidae; Catostomus.
OX   NCBI_TaxID=7971;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT   "Molecular cloning and characterization of white sucker
     preprosomatostatin.";
LR   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF292654; AAK97071.1;
SQ   SEQUENCE 120 AA; 13783 MW; 00828D35263E8805 CRC64;

Query Match      34.6%; Score 196; DB 13; Length 120;
Best Local Similarity 43.3%; Pred. No. 1.3e-11;
Matches 52; Conservative 15; Mismatches 37; Indels 16; Gaps 4;

QY   6 VQCALALLSLALAISSVSA-----APSDAKRLQLLQRSILMAPAGKQELARNTLVLLSEL 60
Db   : | ||| | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
        6 LHCYLALLGLSLVLCRGADSQLPDMDFRHRLLQRARAIGLATQDWTKKDIEELLSQL 65
QY   61 AHVENEAEIELDDMSHGVE---QEDVDLELERAPGV--LAPRRKAGCKNFFWKTFTSC 114
Db   : | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
        66 SLPEIEARE-----NGVSTTGNDLDHLELSAENTQLYPERRKAGCKNFFWKTFTSC 120

RESULT      7
Q90XEO      PRELIMINARY;          PRT;    111 AA.
ID   Q90XEO
AC   Q90XEO;
DT   01-DEC-2001 (TREMBLrel. 19, Created)
DT   01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE   SOMATOSTATIN PRO2.
OS   Acipenser transmontanus (White sturgeon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC   Acipenserinae; Acipenser.
OX   NCBI_TaxID=7904;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Trabucchi M., Tostivint H., Lihrmann I., Sollars C., Vallarino M.,
RA   Does R.M., Vaudry H.;
RT   "Polygenic expression of somatostatin in the sturgeon Acipenser
RT   transmontanus: molecular cloning and distribution of the mRNAs
RT   encoding two somatostatin precursors.";
RL   J. Comp. Neurol. 0:0-(2001).
DR   EMBL; AF395850; AAL13249.1;
SQ   SEQUENCE 111 AA; 12748 MW; 4E27DB90896A9025 CRC64;

Query Match      20.7%; Score 117; DB 13; Length 111;
Best Local Similarity 35.3%; Pred. No. 0.00054;
Matches 41; Conservative 15; Mismatches 48; Indels 12; Gaps 4;

QY   6 VQCALALLSLALAISS--VSAAPSDAKLRLQLLQRSILMAPAGKQELARNTLVLLSELAH 62
Db   : | | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
        1 MQLRARLVSLMLVYSILRVAVLPGERLSVHSNREL-----SKERKEGFKLLSGLLDR 55
QY   63 VENEAEIELDDMSHGVEDVDLEL-ERAPGPVLAP---RERKAGCKNFFWKTFTSC 114

```


Best Local Similarity 81.0%; Pred. No. 0.006; 1; Mismatches 3; Indels 0; Gaps 0;
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 94 LAPRERKAGCKNFFWKFTTSC 114
Db 8 LPARERKAGCKNFFWKFTTSC 28

RESULT 10
Q9PRV0
ID Q9PRV0 PRELIMINARY; PRT; 25 AA.
AC Q9PRV0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SOMATOSTATIN-RELATED PEPTIDE
OS Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguilloidei;
OC Anguillidae; Anguilla.
OX NCBI_TaxID=7937;
N [1]
RP SEQUENCE.
RX MEDLINE=95053622; PubMed=7525832;
RA Uesaka T., Yano K., Yamasaki M., Nagashima K., Ando M.;
RT "Somatostatin-related peptides isolated from the eel gut: effects on
ion and water absorption across the intestine of the seawater eel.";
RL J. Exp. Biol. 188:205-216(1994).
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
SQ SEQUENCE 25 AA; 2860 MW; BFC672143A04A3F5 CRC64;

Query Match 15.4%; Score 87; DB 13; Length 25;
Best Local Similarity 83.3%; Pred. No. 0.076; 1; Mismatches 2; Indels 0; Gaps 0;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 97 RERKAGCKNFFWKFTTSC 114
Db 8 RERKAGCKNFFWKFTTSC 25

RESULT 11
Q9R1P8
ID Q9R1P8 PRELIMINARY; PRT; 105 AA.
AC Q9R1P8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PREPROCORTISTATIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
N [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=99453346; PubMed=10521599;
RA Calbet M., Guadano-Ferraz A., Spier A.D., Maj M., Sutcliffe J.G.,
RA Przewlocki R., de Lecea L.;
RT "Cortistatin and somatostatin mRNAs are differentially regulated in
response to kainate.";
RL Brain Res. Mol. Brain Res. 72:55-64(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Calbet-Murtro M., de Lecea L.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050156; AAD51127.1;
DR MGD; MGI:109538; Cort.
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.

FT CHAIN 88 105 CORTISTATIN.
SQ SEQUENCE 105 AA; 11061 MW; F0DCD40052AB8E95 CRC64;
Query Match 15.0%; Score 85; DB 11; Length 105;
Best Local Similarity 53.6%; Pred. No. 0.64; 2; Mismatches 0; Gaps 0;
Matches 15; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 87 ERAPGVLAPRERKAGCKNFFWKFTTSC 114
Db 77 QERPPQPPHLDKKCKNFFWKFTTSC 104

RESULT 12
Q63803
ID Q63803 PRELIMINARY; PRT; 715 AA.
AC Q63803;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE XLALPHAS PROTEIN.
GN XLAS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95089824; PubMed=7997272;
RA Kehlenbach R.H., Matthey J., Huttner W.B.;
RT "XLas is a new type of G protein.";
RL Nature 372:804-809(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Klemke M., Kehlenbach R.H., Huttner W.B.;
RT "Two overlapping reading frames in a single exon encode interacting
proteins - a novel way of gene usage.";
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-367 FROM N.A.
RC STRAIN=WISTAR;
RA Wang Y.Z., Kehlenbach R.H., Huttner W.B.;
RT "The XL-domain of rat XLas is encoded by a single exon.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; X84047; CAC39211.1;
DR EMBL; AF093569; AAD03032.1;
DR HSSP; P04896; IAZT
DR InterPro; IPR001019; Gprotein_alpha.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR SMART; SM00275; G-alpha; 1.
SQ SEQUENCE 715 AA; 78232 MW; 68F9AAAC18159896 CRC64;

Query Match 14.0%; Score 79.5; DB 11; Length 715;
Best Local Similarity 30.1%; Pred. No. 20;
Matches 40; Conservative 13; Mismatches 55; Indels 25; Gaps 7;

QY 2 LSTRVQCALALSLALAIIS---SVSAAPSDAKLR-----QLQKSLMAPA-----GK 45
Db 171 LPARAAAAAARAAARAAVAAGRSASAPSRALRPPSPPEIQVADPTTPRPAAPRPSAWPDK 230
QY 46 QELARNTL-VELLSELAHVENEFAIEIEMDMSHGVEQEDVDLELRAPGVLAPRERKAG-- 102
Db 231 YERGRSCCRYEAASGICEIESSESEEGATGCFQ--WLLRRNRPG---QPRSHTVGSN 285
QY 103 -CKNFFWKFTTSC 114
Db 286 PVRNFFARAFGSC 298

RESULT 13
O01392

